The Open Orthopaedics Journal

BENTHAM OPEN

CrossMark
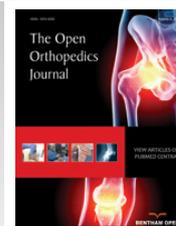
Reliability and Validity of Electro-Goniometric Range of Motion Measurements in Patients with Hand and Wrist Limitations

Siamak Bashardoust Tajali[1,4,*], Joy C. MacDermid[1,2], Ruby Grewal[1,3] and Chris Young[2]

[1]*Department of Health and Rehabilitation (Physical Therapy), The University of Western Ontario, London, Ontario, Canada*

[2]*Hand and Upper Limb Centre, Clinical Research Laboratory, St Joseph's Health Centre, London, Ontario, Canada*

[3]*Hand and Upper Limb Centre, Department of Surgery, St. Joseph's Health Centre, London, Ontario, Canada*

[4]*Department of Physical Therapy, Tehran University of Medical Sciences, Tehran, Iran*

Abstract:

Study Design:

Cross-sectional reliability and validity study.

Purpose:

1. To determine intrarater, interrater and inter instrument reliabilities and validity of two digital electro goniometry to measure active wrist/finger range of motions (ROMs) in patients with limited motion. 2. To determine intrarater and interrater reliabilities of digital goniometry to measure torques of PIP passive flexion of the index finger in patients with limited motion.

Methods:

The study was designed in a randomized block plan on 44 patients (24 women, 20 men) with limited wrist or hand motions. Two experienced raters measured active wrist ROMs, and active and passive index PIP flexion using two digital goniometers. All measures were repeated by one rater 2-5 days after the initial measurements. The reliability measures were analyzed using Intraclass Correlation Coefficients (ICCs) and the construct validity was determined by correlation coefficients analysis between sub measures of scores; patient rated pain and function (PRWE) and quick Disabilities of the Arm, Shoulder and Hand (quick DASH) scores.

Results:

The intrarater, interrater and inter instrument reliabilities were high in most ROM measures (range 0.64-0.97) for both types of electro-goniometers. The 95% limit of agreements and Bland and Altman plots did not show progressive changes. There was a significant difference in force application between the raters when performing passive ROM measures for PIP index, but the same rater produced consistent force. Most of the NK and J-Tech ROM measures were moderately correlated with the patient rated pain and function scores (range 0.32-0.63).

Keywords: Digital goniometry, Hand, Range of Motion, Reliability, Validity, Wrist.

INTRODUCTION

Loss of range of motion (ROM) in the wrist and hand can arise secondary to pain, swelling, muscle weakness, or deformity [1]. Loss of ROM leads to a decrease in grip strength, grasp ability, fine manipulation, and hand function [1]. ROM measurement is considered an important component of hand joint assessment to measure impairment, as well as

---

* Address correspondence to this author at the Hand and Upper Limb Centre, Clinical Research Lab, St. Joseph's Health Centre, 930 Richmond Street, London, Ontario N6A 4V2, Canada; Tel: (519) 646-6100; Ext 61852; Fax: (519) 646-6049; E-mail: sbashar@alumni.uwo.ca

to evaluate the effects of therapeutic interventions [2]. Goniometry is an easy, noninvasive, and inexpensive method of measurement [3] and is considered a precise method to assess movement capability [4].

A number of studies have evaluated the reliability of manual goniometry providing support for current use of goniometry. Flowers *et al*. [5] studied intra and interrater reliability of passive wrist flexion and extension ROM in the patients of eight clinics around the United States. The evaluators (4 therapists in each clinic) randomly measured passive wrist flexion/extension ROM of 141 patients with a plastic manual goniometer and in a blinded design. The authors (who were not the raters) determined that six of the eight clinics had significant differences among the various goniometric techniques. Ellis and Bruton [6] reported about 5˚ difference for intrarater reliability and 7˚ to 9˚ difference for interrater reliability with 95% confidence interval for finger manual goniometry.

Previous studies have provided limited evidence about computerized digital goniometers. Jonsson and Johnson [7] compared ROM measurement accuracy between two types of wrist goniometers; a biaxial single-transducer and a biaxial two-transducer. The research showed that the biaxial single-transducer goniometer had larger errors compared to that of the biaxial two-transducer system. However, neither of these devices is commercially available. Armstrong *et al*. [8] reported intraclass correlation coefficients (ICCs) for forearm flexion/extension and rotation while reporting intrarater, interrater and inter instruments reliability across 5 raters and 3 types of goniometers: a universal standard, an NK computerized goniometers, and a mechanical rotation measuring device. The reliability of the elbow flexion/extension and pronation/supination were moderate to high across different occasions or raters. Rotation measurements tended to have larger errors (SEM) that did elbow flexion/extension measures examined within the same study. However, the bias in three instruments was relatively small, but inter instrument related measurement error was substantial. The researchers also identified that reliable ROM measurements of elbow flexion/extension and forearm pronation/supination were obtainable regardless of the level of experience when the raters used a standard measurement method. The NK Hand assessment system goniometers although reliable are no longer supported commercially, so clinicians who wish to adopt this approach would need to know the reliability of commercially available devices.

Jonsson *et al*. [9] studied the accuracy and feasibility of using an electrogoniometer for measuring simple thumb movements in healthy subjects. The researchers compared the results of eight positions for thumb flexion/extension and abduction/adduction between digital and manual goniometers and indicated that the only significant difference was found between the goniometers when the thumb was in full flexion. The researchers identified that electrogoniometric measurement errors were lower than 5˚ for the thumb ROM measures in comparison to manual goniometry. The researchers recommended the use of the electrogoniometers for studying thumb based activities, since it could provide quantitative information on thumb movements during thumb intensive tasks.

A reliable ROM measurement helps clinicians make a treatment plan based on accurate measurement of motion impairments. Although manual goniometers have a stable position in hand therapy practice, the use of computerized tools is expected to increase over time as the costs become lower; and as computers become integrated in other aspects of practice. The digital electrogoniometric devices (such as NK and J-Tech) potentially offer mechanical precision and reduced rater reading errors; and thus may enhance the accuracy of assessment of hand joint ROM, mobility and severity of impairment. The NK device has advantages; in that we already know it is precise; while the J-tech has advantages in that is commercially available as part of a complete hand assessment system designed for clinical practice.

The NK torque-motion goniometer allows assessment of torque applied when a given joint motion is measured. Torque values cannot be measured by traditional manual goniometers; unless extra instrumentation is applied. It has been suggested that torque ROM measurements can inform our understanding of the extensibility of soft tissues limiting ROM; and thus could contribute to selecting and evaluating interventions [10, 11]. For instance, tendon transfer surgery may be indicated in patients with flexion contracture associated with median/ulnar nerve palsy based on steep torque angle curves that indicate a lack of potential for tissue elongation through conservative measures [10]. A further purpose of torque goniometry is to understand the force applied while assessing ROM, since it is assumed that this might contribute to differences in motion estimates obtained by different raters. Patients are often measured repeatedly during recovery. Thus, it is important to know how comparable these measures are likely to be.

The primary purpose of this study was to determine the intrarater, interrater and inter instrument reliability and criterion related validity of wrist and PIP index finger ROM measures using two digital electrogoniometric devices in patients with limited wrist and/or hand motion. The null hypothesis was that there were no significant differences in wrist and PIP index measures between occasions, across raters or instruments using two digital electro goniometric

devices. The secondary purpose was to assess whether the torque applied during ROM measurements varied across different raters; using PIP passive flexion of the PIP index finger as the construct.

MATERIALS AND METHODOLOGY

Study Design

The study was designed as a cross-sectional reliability and validity study, so that the reliability of two digital electrogoniometry instruments was assessed between two occasions, across two raters and between two instruments.

Participants

Patients with limited wrist and/or hand motion who met eligibility criteria and consented to participation were enrolled in the study. Participants were included if they were 19 years of age or older and had limited wrist and/or hand motion 8 to 12 weeks following a musculoskeletal disorder. They also must have been able to speak and understand English and learn simple instructions. Patients were excluded from the study if they were under 19 years old, unable to follow study instructions, had an acute infection or open wound, a history of neurological or rheumatologic conditions, bilateral hand disorders or combined arm/shoulder or multiple disabled joints.

Forty nine patients participated in the study, and a written consent form was obtained before measurement. All participants were outpatients of the Hand and Upper Extremity Center (HULC) at St. Joseph Health Care Center in London Ontario. The participants were recruited and measured within the initial eight to twenty four weeks of their injury. All participants completed a brief survey including demographic data (age, gender, affected side, medical history, *etc*) before data collection. The study was reviewed by the university and hospital academic and ethical boards and was approved before starting data collection.
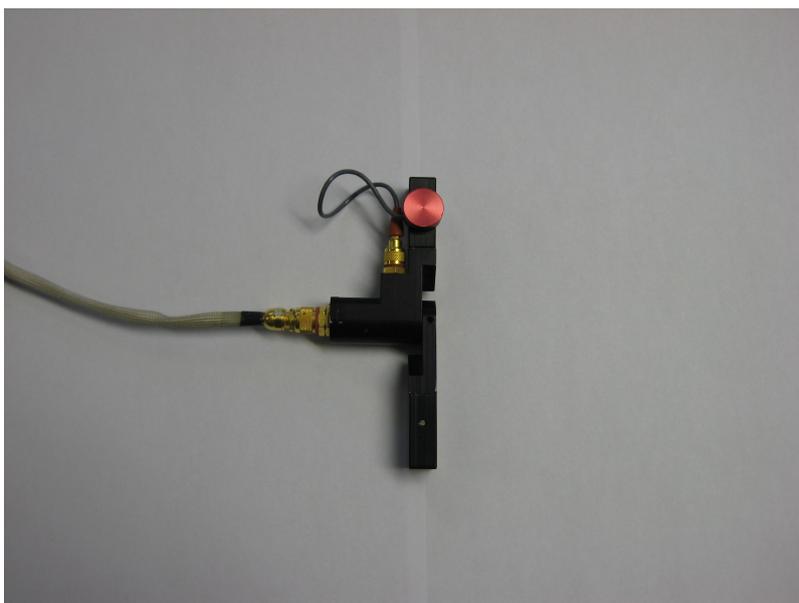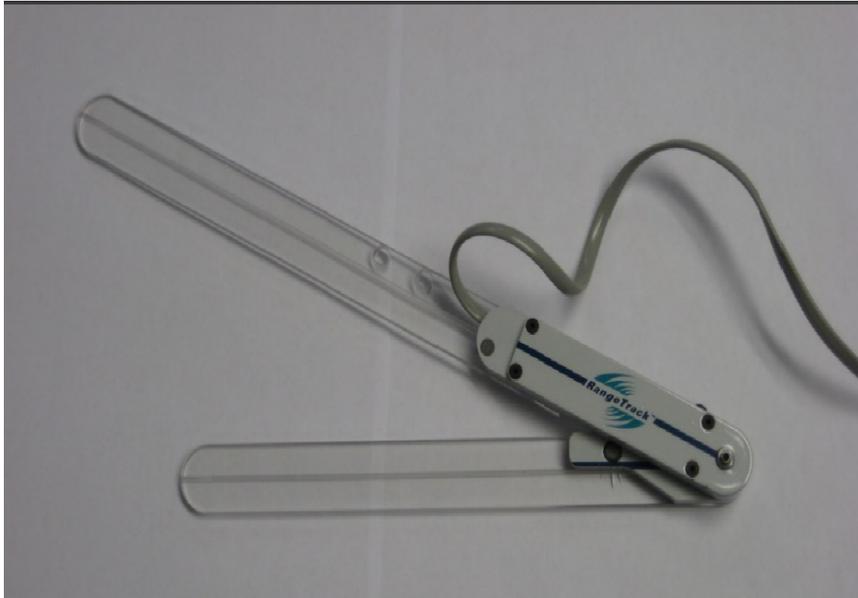

Fig. (1). NK digital goniometer.

Raters and Instruments

Two raters obtained the measurements in two different sessions. One rater was a PhD physical therapist and the other a kinesiologist. Both raters had clinical experiences measuring range of motion for more than five years. The raters used the NK Hand Assessment Laboratory joint motion (NK Biotechnical Engineering Company, Minneapolis), and the J-Tech electrogoniometer (JTech Medical, Salt Lake City, UT) and their associated software for ROM measurements. The NK and J-Tech are two instruments which can be used to assess hand joint ROM, mobility and severity of impairment (Figs. **1** and **2**). Data collection was performed with standard computer software sensitive with a foot switch, so that the rater's hands were free to adjust the goniometric alignment. Active ROM of the wrist motion (flexion and extension, radial/ulnar deviation, pronation and supination), and active and passive ROM of proximal inter phalangeal (PIP) joint of the index finger (flexion) were measured for each participant by both NK and J-Tech Hand

electrogoniometers. There was a self calibrating device in both electronic measurement instruments so that the raters could calibrate both instruments prior to the study and before each measurement. The lengths of the arms were equal in NK (2 inches), while the lengths of the short and long arms were 7.5 and 10.5 inches in J-Tech. The NK digital instrument had a specific gauge and a digital force transducer which could be used to measure the amount of passive force applied for the hand ROM measurements. Patients were asked if they were relaxed and comfortable before the measurements were taken. Physical disturbances such as fatigue, sleepiness, dizziness, alcohol or usage of drug that might affect the measurement results were asked before the measurements.

Fig. (2). J-Tech digital goniometer.

Patient Positioning: Three positions were used for different ROM measurements. To measure wrist flexion/extension, ulnar/radial deviation, and index finger flexion, each participant sat in front of a hand assessment table with their elbow placed on the table. The elbow was held in $110°$ - $120°$ of flexion for wrist flexion/extension and PIP flexion measurements, and was held at $90°$ flexion for the measurements of radial/ulnar deviation. To measure wrist pronation/supination, each participant stood in front of the assessment table and kept her/his arm close to the body and the elbow was positioned at $90°$ of flexion. The forearm was in neutral position for all measurements [12, 13].

Landmarks: Established reliable landmarks were used for goniometry [12]. The raters reviewed and agreed the landmarks before they started the ROM measurements. The following landmarks were used to measure ROMs:

1. *Wrist Flexion:* The stationary arm was aligned on the dorsal midline of the forearm, the movable arm on dorsal surface of third metacarpal, and the center fulcrum over the capitates on the dorsal aspect of wrist (Fig. **3**). *Wrist extension:* The stationary arm was aligned on the palmar midline of forearm, the movable arm over palmar midline surface of third metacarpal, and the center fulcrum on the palmar surface of the wrist at the level of the capitate [13].
2. *Radial/Ulnar Deviation:* The stationary arm was aligned on midline of dorsal surface of forearm, the movable arm dorsally over midline of third metacarpal, and the center fulcrum on capitate [12] (Fig. **4**).
3. *Pronation:* The stationary arm was at the dorsal aspect of the wrist paralleled to anterior longitudinal midline of humerus, the moveable arm on the widest dorsal area of the wrist proximal to the styloid processes of radius and ulna, the center fulcrum on lateral and proximal aspect of ulnar styloid process (Fig. **5**).
4. *Supination:* The stationary arm was at ventral aspect of the wrist parallel to anterior longitudinal midline of humerus, the moveable arm on volar surface of the wrist at level of ulnar styloid processes, the center fulcrum on volar surface of the wrist in line with ulnar styloid process [12].
5. *Index PIP Flexion:* The stationary arm was aligned dorsally over proximal and the moveable arm over middle phalanxes, the center fulcrum dorsally over PIP joint [12].

Fig. (3). Flexion measurement by J-Tech.

Fig. (4). Ulnar deviation measurement by NK.

Data was collected on 2 separate days with 2-5 days between sessions. The raters used a random number generator program to randomize both raters and instruments for each participant (random.org). On the first occasion, the ROM measurements were taken in random order by rater one or rater two and for the NK or J-Tech goniometer. After a short period of rest (5 minutes), the second rater performed the similar ROM measures for wrist and index finger motions. After a longer period of rest (10 minutes), the ROM measurements were repeated in a similar way by the other digital goniometer (NK or J-Tech). On he second occasion, 2-5 days later, the first author repeated digital goniometry ROM measures with both instruments in random order (Fig. **6**).
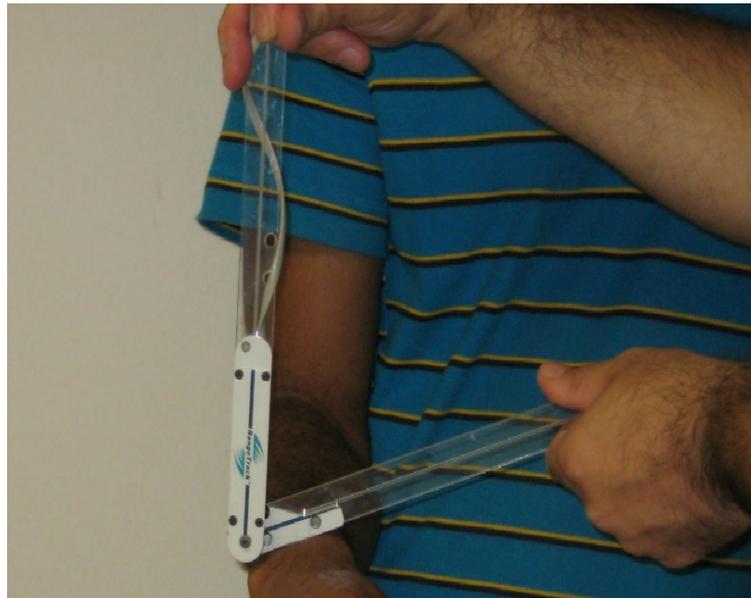
Fig. (5). Pronation measurement by J-Tech.

In each measurement session, the participants were measured in a sitting position. The raters asked the participant to actively perform a maximum movement and were given a brief instruction and 2 practice trials prior to recording scores. The mean of three repetitions were taken as data for each ROM measure. Following the active ROM measurements, passive ROM of PIP index flexion was taken only with the NK instrument. For the torque measurements of passive flexion of the PIP index, the raters manually held the metacarpophalangeal (MCP) joint at neutral position throughout the testing procedure. Then, the raters applied a flexion force perpendicular to the middle phalanx at the dorsal surface over the PIP index and at the ending range of active flexion of the PIP index. The transducer recorded each force measurement and with the average of three torque measures considered as the torque value for passive PIP flexion of index finger in each session.

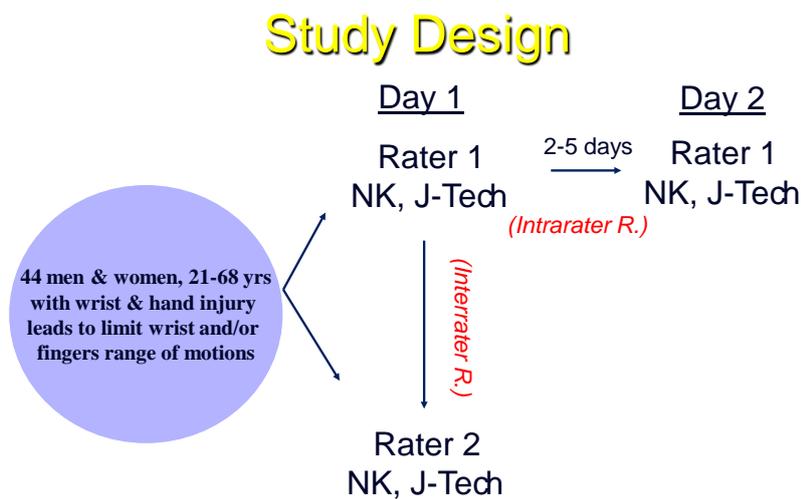**Fig. (6).** Diagram of the study design.

A Patient Rated Wrist Evaluation (PRWE) questionnaire and a standard short version of the Disability of the Arm, Shoulder and Hand (Quick DASH) questionnaire, were completed by the participant before or after the first session of the measurements. Data was recorded by the relevant software in each instrument and transferred to a data collection form by the raters.

Statistical Analysis

Data was analyzed by SPSS version 19 (SPSS Inc., Chicago). Descriptive statistics were reported based on means ± SD. Tests of difference and reliability coefficients were calculated to compare the data between different occasions, raters and instruments. Repeated measures analyses of variances (RM-ANOVA) were used to determine similarity of the ROM results obtained on different occasions or across raters. If the results were statistically significant, multiple comparison post hoc Tukey Honestly test were performed to determine which means were different from the others. The Tukey Honestly post hoc test is one of the most conservative multiple comparison designs [14]. A factorial ANOVA was used to identify the interaction effects among the ROM results (dependant variable) across the raters and the electro goniometers (fixed variables). This analysis informs whether or not these two different measurement techniques can be used interchangeably [5].

Intraclass correlation coefficients model (2,1) ($ICC_{2,1}$) and their associated 95% confidence interval (CI) were calculated [15, 16] to compare the scores of each measurement across occasions in same rater (intrarater reliability), between the raters (interrater reliability), and between the instruments for each rater (inter instrument reliability). The $ICC_{2,1}$ was used to represent the scores by two raters or instruments and a single measure was taken for each of them [17]. We used the mean results of three repetitions for each measurement per session. The ICC values of each rater in first day of measurements were used for intrarater reliability analyses, and the ICC values of rater one in first and second days of measurements were used for interrater reliability analyses. The cut-off scales of ICC >0.75, 0.40-0.75, and <0,40 were chosen as an indication for high, moderate, and low reliability, respectively [18]. The Standard Error of Measurement (SEM) was calculated to identify absolute reliability of the measures and estimated the measurement error in a set of repeated scores [15]. The SEM is calculated by the equation $SEM = SD \times \sqrt{1-r}$ [17, 19]. The Minimum Detectable Change (MDC) was calculated to define the smallest amount of change needed to be certain that a real change was occurring beyond a measurement error [20]. The MDC was calculated with 90% and 95% confidence interval using the specific equation ($MDC\% = z_{(df, \alpha)} \times SEM \times \sqrt{2}$)[13].

The agreement parameters show the size of the measurement errors [21]. We calculated 95% limits of agreement (LoA) and constructed Bland and Altman plots to account for potential systematic bias between the raters or instruments. Bland and Altman plots are the more commonly seen description that graphically demonstrate the agreement between these measures [21]. The LoA was calculated based on the equation LoA= Mean difference + 1.96 × Standard Deviation (SD). The mean differences describe any systematic difference (bias) between measurements. The limits of agreement defines the range in which repeated measurements might be expected to vary with 95% confidence [22]. The association between motion measures and PRWE or DASH was described by Pearson's r correlation coefficients. Pearson correlation r <0.40, between 0.40-0.75, >0.75 were considered as low, moderate and high [23]. The alpha was 0.05 and the results were considered significant if $p < 0.05$.

RESULTS

Three participants were excluded because they presented with rheumatologic (one patient) or neurologic-stroke (two patients) conditions that might affect their wrist and hand motion. Two participants dropped out and did not attend the second occasion. In total, 44 participants completed the study (24 women and 20 men; 55% *vs*. 45%), with an age range between 21 to 68 years old (52.50 ± 12.92). Twenty one participants (47.7%) had an injury on their dominant hand, while twenty three (52.3%) had an injury on their non-dominant hand. A chi-square test of independence showed that there were no significant difference between the proportions of dominant and non-dominant injured sides [$x^2_{(1)}$ 0.72, NS]. The participants' height and weight were 172 ± 12 cm and 77 ± 21 Kg. The initial diagnosis of participants were: 32 patients (73%) distal radius fracture, 6 (14%) carpal tunnel syndrome, 3 (7%) scaphoid fracture, 2 (4%) finger fracture, 1 (2%) metacarpal fracture.

The summary of means ± SDs for the occasions and raters in both instruments and ANOVA statistical analysis to compare the ROM measures in different occasions were not substantially different between the raters for each goniometer. The raters did not demonstrate consistent use of force when performing passive ROM measures for PIP index flexion as the Tukey post hoc test showed that there were significant differences in torque applied by the two raters during passive ROM measures for index PIP flexion ($F_{(1, 42)}$ 44.17, p 0.00, q 12.60) (Supplementary Table **1**).

The factorial ANOVA for main effects (rater and instrument) and interaction effects (rater × instrument) showed that there were no interaction effects through the outcome measures (Supplementary Table **2**). The raters did not affect results of ROM measures; however, type of the instrument affected the results of ROM measures for wrist extension

($F_{(1, 42)}$ 5.09, p 0.02), ulnar deviation ($F_{(1, 42)}$ 5.96, p 0.02), and pronation ($F_{(1, 42)}$ 8.80, p 0.03) (Supplementary Tables **1** and **2**).

Table 1. Intrarater (test-retest) reliability values for NK and J-Tech digital goniometers.

Measure/Reliability	Instrument	ICC	95%CI	SEM	MDC$_{90}$	MDC$_{90}$
Wrist Flex	NK	0.97	0.95-0.98	1.98	4.62	5.49
	J-Tech	0.95	0.91-0.97	2.59	6.04	7.18
Wrist EXT.	NK	0.95	0.91-0.97	2.06	4.81	5.71
	J-Tech	0.94	0.72-0.94	2.47	5.76	6.85
Wrist Radial Dev.	NK	0.96	0.92-0.98	0.96	2.24	2.66
	J-Tech	0.93	0.88-0.96	1.26	3.04	3.49
Wrist Ulnar Dev.	NK	0.91	0.85-0.95	1.98	4.62	5.49
	J-Tech	0.93	0.89-0.96	2.06	4.81	5.71
Wrist pron.	NK	0.89	0.80-0.94	2.10	4.90	5.82
	J-Tech	0.86	0.77-0.92	2.24	5.23	6.21
Wrist Sup.	NK	0.94	0.85-0.96	3.20	5.13	6.10
	J-Tech	0.95	0.90-0.97	1.95	4.55	5.40
Act.PIP Index Flex.	NK	0.93	0.89-0.97	1.78	4.15	4.93
	J-Tech	0.91	0.84-0.95	1.93	4.50	5.35
Pas. PIP Index Flex.	NK	0.91	0.85-0.95	1.90	4.43	5.27
Torque	NK	0.71	0.44-0.89	7.57	17.66	20.98

Note: ICC = Intraclass Correlation Coefficient; CI = Confidence Interval; SEM = Standard Error of Measurement; MDC$_{90}$ = Minimum Detectable Change associated with 90% CI; MDC$_{95}$ = Minimum Detectable Change associated with 95% CI.

Table 2. Interrater (between raters) reliability values for NK and J-Tech digital goniometers.

Measure/Reliability	Instrument	ICC	95%CI	SEM	MDC$_{90}$	MDC$_{95}$
Wrist Flex	NK	0.97	0.95-0.98	3.64	8.49	10.09
	J-Tech	0.87	0.88-0.97	3.48	8.12	11.86
Wrist EXT.	NK	0.95	0.91-0.97	2.06	4.81	9.65
	J-Tech	0.93	0.90-0.97	2.82	6.58	7.82
Wrist Radial Dev.	NK	0.84	0.73-0.91	2.16	5.04	5.99
	J-Tech	0.87	0.72-0.93	1.94	4.53	5.76
Wrist Ulnar Dev.	NK	0.82	0.51-0.92	2.60	6.07	7.17
	J-Tech	0.93	0.89-0.96	2.06	4.81	5.38
Wrist pron.	NK	0.83	0.71-0.90	3.11	7.26	8.62
	J-Tech	0.79	0.69-0.89	3.02	7.05	8.37
Wrist Sup.	NK	0.87	0.73-0.93	3.78	8.82	10.48
	J-Tech	0.84	0.49-0.94	3.96	9.24	10.98
Act.PIP Index Flex.	NK	0.85	0.74-0.91	2.81	6.56	7.79
	J-Tech	0.81	0.69-0.90	2.75	6.42	7.62
Pas. PIP Index Flex.	NK	0.83	0.71-0.90	2.63	6.14	7.29
Torque	NK	0.16	0.08-0.23	9.83	22.94	27.25

Note: ICC = Intraclass Correlation Coefficient; CI = Confidence Interval; SEM = Standard Error of Measurement; MDC$_{90}$ = Minimum Detectable Change associated with 90% CI; MDC$_{95}$ =Minimum Detectable Change associated with 95% CI.

The ICC values for intrarater reliability (test-retest) were excellent for most wrist ROM measures (flexion, extension, radial deviation, ulnar deviation, supination), and PIP index flexion measures by both instruments (ICC ranges 0.91-0.97). The intrarater reliability was also high for wrist pronation measures for both NK (ICC 0.89) and J-Tech (ICC 0.86). The highest intrarater reliability values were in wrist flexion ROM measures by both the NK (ICC 0.97) and J-Tech (ICC 0.95). The lowest intrarater values were measured in wrist pronation measures by the NK (ICC 0.89) and also J-Tech (ICC 0.86) (Table **1**). The ICC values for interrater reliability were high for active and passive ROM measures (ICC ranges 0.79-0.93). The highest interrater reliability values were in wrist flexion ROM measures by the NK (ICC 0.91) and wrist extension ROM measures by the J-Tech (ICC 0.93). The lowest interrater reliability values

referred to ulnar deviation ROM measures by the NK (ICC 0.82) and pronation ROM measures by the J-Tech (ICC 0.79). The ICC values for inter instrument reliability were high in all wrist ROM measures (ICC ranges 0.77-0.96), with the exception of radial deviation (ICCs 0.64 and 0.70 for the raters one and two, respectively). The reliability coefficients for torques in passive index flexion were moderate in different occasions by rater one (ICC 0.71) and low across the raters (ICC 0.16) (Tables **2** and **3**).

Table 3. Inter instruments reliability values for NK and J-Tech digital goniometers.

Measure/Reliability	Instrument	ICC	95%CI	SEM	MDC	MDC
Wrist Flex	Rater 1	0.96	0.92-0.98	2.37	8.49	10.09
	Rater 2	0.94	0.89-0.97	2.93	6.84	8.12
Wrist Ext.	Rater 1	0.89	0.81-0.94	3.07	7.16	8.51
	Rater 2	0.93	0.88-0..96	2.74	6.39	7.59
Wrist Radial Dev.	Rater 1	0.64	0.42-0.79	2.79	6.51	7.73
	Rater 2	0.70	0.52-0.83	2.92	6.81	8.10
Wrist Ulnar Dev	Rater 1	0.86	0.77-0.92	2.49	5.81	6.90
	Rater 2	0.87	0.78-0.93	3.11	7.26	8.62
Wrist Pron	Rater 1	0.77	0.62-0.87	3.17	7.40	8.79
	Rater 2	0.78	0.64-0.88	3.86	9.01	10.70
Wrist Sup	Rater 1	0.94	0.88-0.96	2.75	6.42	7.62
	Rater 2	0.90	0.83-0.94	3.25	7.58	9.01
Act. PIP Index Flex	Rater 1	0.89	0.82-0.94	2.41	5.62	6.68
	Rater 2	0.79	0.64-0.88	3.05	7.12	8.45

Note: ICC = Intraclass Correlation Coefficient; CI = Confidence Interval; SEM = Standard Error of Measurement; MDC_{90} = Minimum Detectable Change associated with 90% CI; MDC_{95} =Minimum Detectable Change associated with 95% CI.

The standard error of measurement (SEM) indicated higher inter-rater error (1.94-9.83) than intrarater (0.96-7.57). The 90 and 95 minimum detectable change (MDC) calculation indicated that 90% (95%) of participants provided less than 4.6˚ (5.5˚) variation when the wrist flexion ROM was measured by one rater in different occasions, and less than 8.5˚ (10.1˚) variation when the wrist flexion ROM was measured by two raters in same occasion (NK goniometer). The SEM and MDC scores between instruments were similar to that between raters' measurements (Tables **1**-**3**).

The highest level of agreement between the raters was found for ulnar and radial deviation ROM measures for both instruments (LoA -4.3 to 10.7), while the torque measures of passive PIP index ROM flexion by NK goniometer had the widest limits of agreement across the raters (LoA -66.3 to 14.5) (Table **4**). The most precise limits of agreement between the instruments was in active PIP index flexion for both raters (LoA -6.71 to 4.81 for rater one; -7.68 to 8.36 for rater two), while the lowest level of agreement between the instruments for rater one was in wrist extension (LoA -12.64 to 5.40), and for rater two was in pronation (LoA -11.83 to 7.03) (Table **4**). The Bland - Altman plots and scatter of mean differences between measurements (raters or instruments) did not show progressive changes across the range of ROM measures (no heteroscedasticity) [24] (Figs. **7**-**9**).

**Table 4. Limit of agreement analysis for the ROM measures across raters or goniometers.**

| Measure			LoA (across raters)	LoA (across instruments)
	NK	J-Tech	Rater 1	Rater 2
Wrist Flexion	-8.18 to 11.66	-6.32 to 12.22	-9.92 to 3.40	- 10.04 to 5.96
Wrist Extension.	- 9.98 to 5.78	- 7.01 to 5.19	- 10.64 to 5.40	-9.94 to 5.08
Wrist Radial Dev.	- 5.39 to 5.55	- 6.19 to 6.97	-8.91 to 7.17	-9.12 to 7.00
Wrist Ulnar Dev.	- 4.33 to 10.69	- 5.29 to 9.17	- 9.37 to 5.65	- 10.35 to 5.15
Wrist Pronation.	- 7.22 to 7.64	- 6.35 to 7.25	- 10.50 to 5.22	- 10.43 to 7.03
Wrist Supination	- 6.62 to 11.36	- 6.35 to 7.25	- 8.07 to 4.47	- 9.95 to 7.15
Act. PIP Index Flex	- 7.83 to 6.99	- 6.63 to 8.39	- 6.71 to 4.81	- 7.68 to 7.36
Pas. PIP Index Flex	- 7.11to 6.37	--	--	--
Torque	- 66.32 to 14.54	--	- -	--

Note: LoA = 95% Limit of Agreement

The relationship between ROM measures and patient rated self-reported pain and function indicated a low to moderate relationship ranging from 0.32 to 0.63. Both the NK and the J-Tech were moderately correlated with self-reported disability of additional error margin was found between (Table **5**). The Pearson's r correlation coefficient between the functional outcome measures (Quick DASH and PRWE) were high (r= 0.94).

(A)

(B)
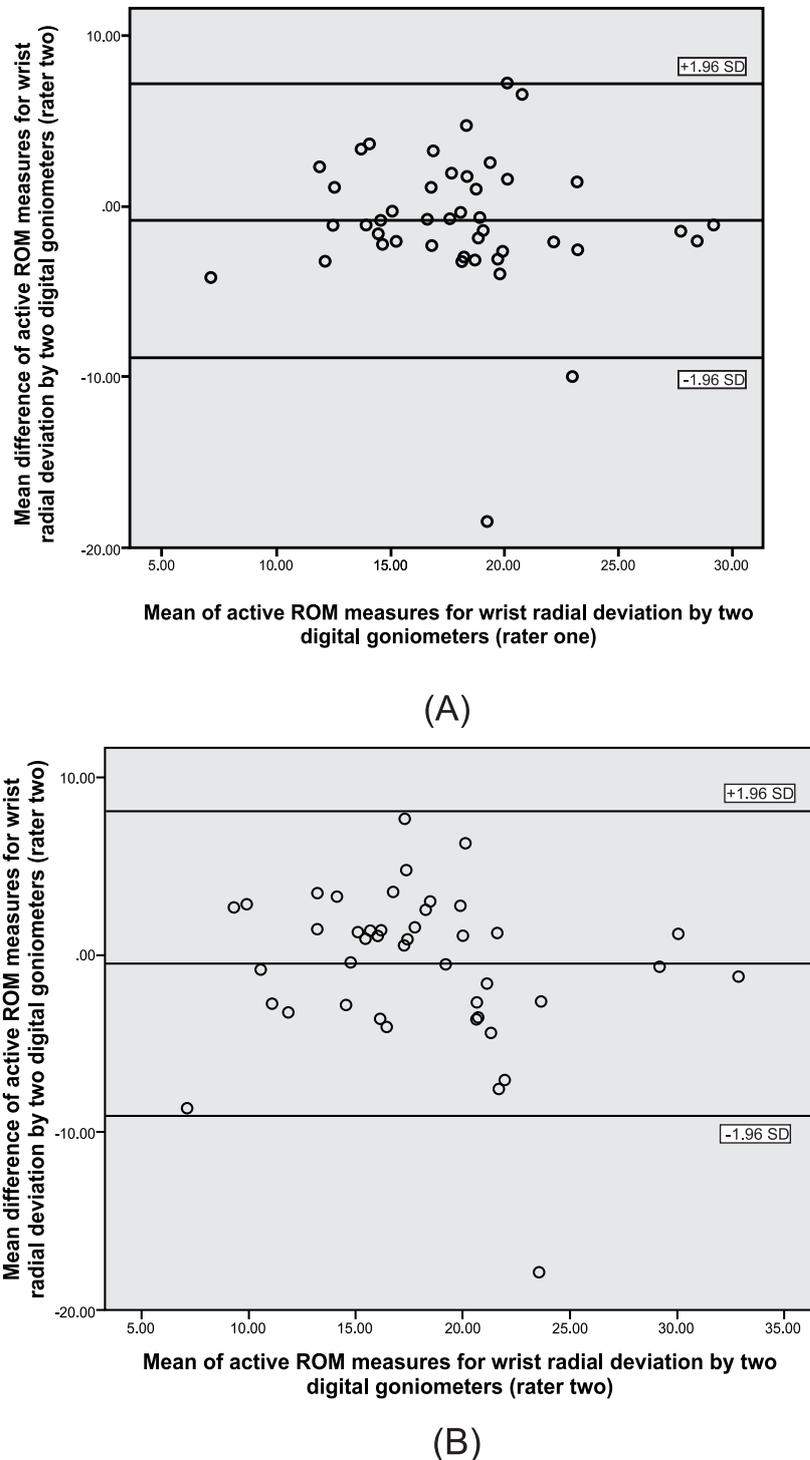
Fig. (7). Bland and Altman plots of mean differences (vertical axis) versus means (horizontal axis) of radial deviation ROM measures by two digital goniometers: (**A**) rater one, (**B**) rater two. The middle line shows the mean difference between measures taken with two digital goniometers (NK-JTech). *The lines above and below mean difference represent range of measurement error with 95% confidence interval (data in degrees).*
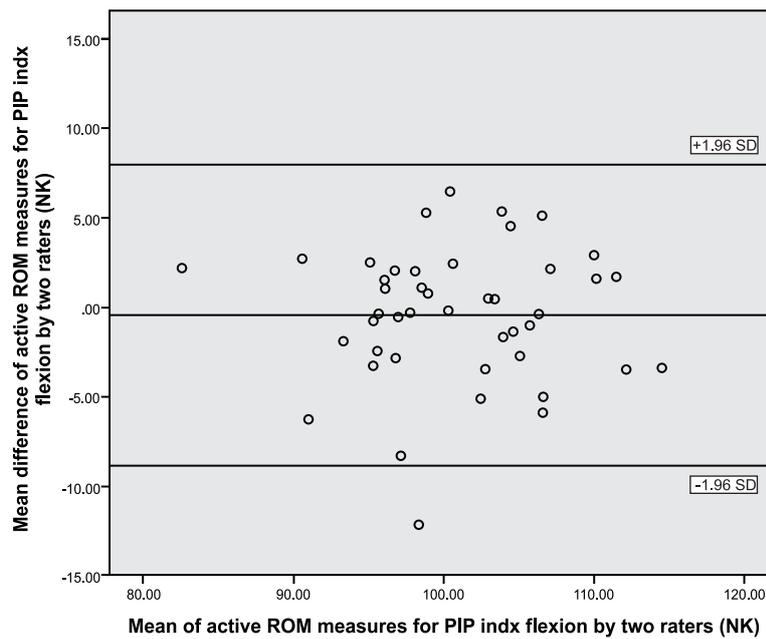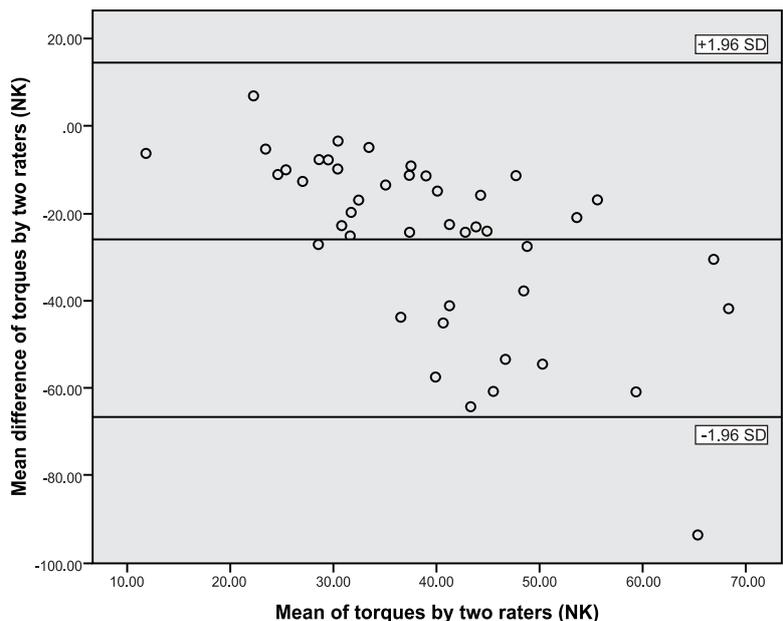
(A)
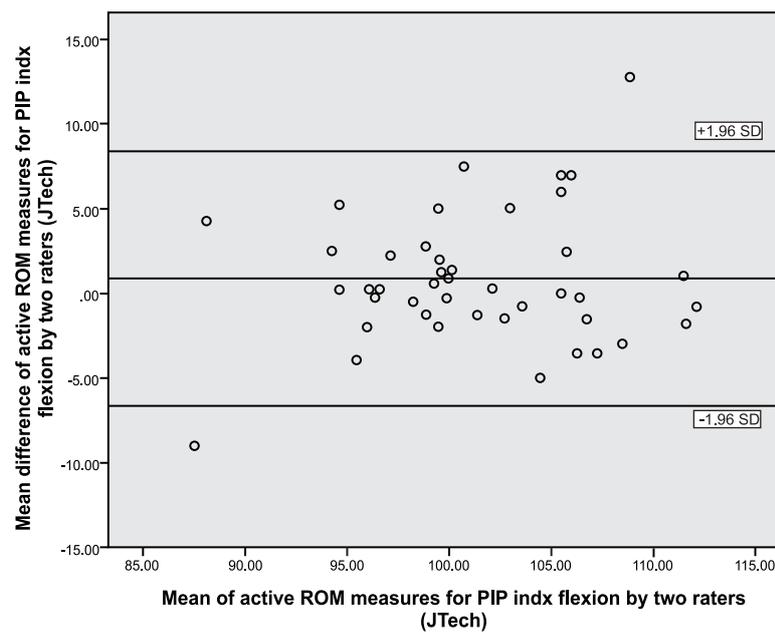
(B)

Fig. (8). Bland and Altman plots of mean differences (vertical axis) versus means (horizontal axis) of active ROM measures for PIP index flexion by two raters; (**A**) NK goniometer, (**B**) J-Tech goniometer. The middle line shows the mean difference between measures taken with two raters in each instrument. *The lines above and below mean difference represent range of measurement error with 95% confidence interval (data in degrees).*

Fig. (9). Bland and Altman plot of mean differences (vertical axis) versus means (horizontal axis) of torques of passive PIP index flexion ROMs by two raters (NK instrument). The mean difference between measures taken with two raters is noticeable. *The lines above and below mean difference represent range of measurement error with 95% confidence interval (data in degrees).*

Table 5. Pearson's r correlation coefficient between the ROM measurements of both digital goniometers (NK and JTech) and patient rated pain and function scores.

Measure			*r correlations with PRWE*				*r correlations with quick DASH*		
	Rater 1		**Rater 2**		**Rater 1**		**Rater 2**		
	NK	**J-Tech**	**NK**	**J-Tech**	**NK**	**J-Tech**	**NK**	**J-Tech**	
Wrist Flexion	-0.44	-0.41	- 0.48	-0.44	-0.41	-0.45	- 0.45	-0.43	
Wrist Extension.	- 0.63	- 0.55	- 0.51	- 0.55	-0.48	-0.48	- 0.45	-0.44	
Wrist Radial Dev.	- 0.44	- 0.41	- 0.48	- 0.41	-0.40	-0.47	- 0.44	-0.36	
Wrist Ulnar Dev.	- 0.39	- 0.44	- 0.39	- 0.35	-0.39	-0.45	- 0.32	-0.35	
Wrist Pronation.	- 0.38	- 0.35	- 0.34	- 0.37	-0.34	-0.38	- 0.38	-0.46	
Wrist Supination	- 0.52	- 0.48	- 0.46	- 0.46	-0.48	-0.46	- 0.52	-0.55	
Act. PIP Index Flex	- 0.46	- 0.50	- 0.47	- 0.48	-0.55	-0.50	- 0.45	-0.50	
Pas. PIP Index Flex	- 0.46	--	0.42	- -	-0.46	--	- 0.48	--	
Torque	-0.18	- -	-0.05	- -	-0.29	--	- 0.16	--	
(Pas. PIP Index Flex)									

Note: PRWE = Patient Rated Wrist Evaluation; quick DASH = short version of the Disability of the Arm, Shoulder and Hand.

Bold = Significant at P <0.05

DISCUSSION

This study demonstrated that reliable measurements of wrist and finger motion are obtainable on different occasions and across different raters with two different computerized goniometers; despite the fact that different raters do not provide consistent pressure when taking passive flexion ROM measurements for PIP index finger. As we expected, the ICCs were slightly higher when the ROM measures obtained by same rater compared than when the ROM measures obtained by two raters. The fact that raters tend to use more consistent force than occurs between raters, suggests that the application force may make a small contribution to lower group-level reliability in PIP index finger ROM measures. However, since the reliability coefficients were, this did not important differences in the measurements obtained. This may be because both raters were able to achieve end range; and the application of extra force did not make an appreciable change. Since the PIPJ is a joint with a hard end feel, it is not clear that this finding will be transferable to other joints with a soft-tissue end feel like elbow flexion.

There are a limited number of studies that measured reliability and validity of wrist and finger ROM measurements. These studies mostly focused on healthy people [6, 7, 9] or patients with normal ROM[5] who were measured by different therapists or occasions. Some studies have addressed digital goniometry, but not included the wrist [8, 9]. Thus study adds additional information on goniometry as we examined rater and instrument effects, and included construct validity comparing to patient reported function.

Our findings of high reliability are in agreement with previous studies that use electrogoniometer for elbow pronation/supination[7] and healthy thumb ROMs measures [9], and also the studies that use manual goniometry for wrist ROM measures. [5 - 7] The precision of measurement compared favorably with what has been reported for manual goniometry suggesting that some small advantages in precision may be obtained by the use of computerized goniometry. Potential reductions in random error with the electrogoniometers might occur to do mechanical precision; use of the footswitch which allows focusing on placement, the computer is not subject to human error in reading the measurement from the device. Plastic goniometers are not calibrated; whereas computerized goniometers are suggesting that variances between different devices would be lessened. Radial deviation ROM measure was the only measure that did not demonstrate high reliability. Possible reasons including difficulty in precise land marking for this movement; and the relatively small ROM measures of the radial deviation should be considered.

The high level of agreements was satisfactory for the raters or instruments in most of the measures; however, our analyses indicate differences in application of force between raters when performing goniometry of PIP index passive flexion. This was demonstrated by low rater agreement; and the significant differences in force application. Since there was only one pair of raters, it is risky to generalize the reasons for differences in force application. It is necessary to clarify that the physical therapist had more experience with patients' measures; and the kinesiologist had more experience to measure ROMs in healthy people. As a result, type of experiences applied by the raters may have affected the raters force application. Since both raters were healthy male, gender was not an intervening factor in the measurements. The random error might be affected by the local pain following inconsistent positioning of the MCP joint and/or following differences in force application by therapists. Standardized methods are the main method for reducing random error. Since on average, $2 – 3°$ of additional error margin was found between intra *versus* inter rater SEMs variations between people on application technique contribute small additional error considerations to scores obtained. The ICCs indicated that provided force used by the same rater was moderately consistent; further supporting this hypothesis. Differences between the digital goniometers for wrist extension, ulnar deviation and pronation ROM measures may have been differences in landmarking by different goniometers. The ICC is a ratio and depends to the range of variation in the sample [25] and lack of spectrum in our sample in some movements may have reduced ICCs.The intrarater and interrater ICCs for ROM measures of pronation and supination had slightly lower reliability and this is consistent with findings of other studies [8].

The Bland and Altman plots across mean differences between measurements (raters or instruments) did not show systematic differences in error across the spectrum of motion. That is there is no heteroscedasticity in ROM measures across the raters or instruments, which means the variance of the error terms does not differ across observations [24]. Systematic errors are sources of bias that should be examined in reliability studies since they can explain differences in scores based on spectrum, occasion or rater. The Bland and Altman technique is better at detecting such differences since it allows for more direct observation of errors than summary statistics like the ICC or SEM.

In recent years, evaluating outcomes of wrist and hand injuries has focused on measuring the patient perspective on pain and disability (PRWE or DASH) [26 - 30]. This study indicated the importance of ROM measurements as contributors to upper extremity-related disability. The relationship between the DASH and PRWE scores and ROM was moderate across the two scales which is consistent with previous findings [26]. These moderate correlations are in concordance with the findings has been found in a variety of other musculoskeletal conditions [30, 32]. Since no single impairment can be expected to fully explain disability, the correlations observed support the current approach to include ROM measures as one aspect of upper extremity assessment. ROM was slightly more strongly related to the PRWE compared to QuickDASH; which may be related to the fact that the PRWE is a wrist-specific scale. Karnezis and Fragkiadakis [28] reported that grip strength could be considered as a predictor for patient-rated pain and function (PRWE), but arcs of wrist flexion/extension and forearm rotation did not. Adams *et al*. [26] reported significant relationship between patient-rated function (DASH and PRWE) and ROMs limitations. MacDermid *et al*. [31] identified a correlation between grip strength, ROM, dexterity (objective variables) and patient-rated pain and function (PRWE) after distal radius fracture, but they also reported these outcome measures could not be considered strong predictors for pain and disability [32].

Study Limitations and Future Steps

Although this study used research design principles like randomization and verification of landmarks; the study also had limitations. To minimize overall subject burden, we only measured finger flexion of one joint. We cannot be confident that this one finger flexion measurement reliability is representative of the reliability of all digits. We examined reliability and construct validity, but did not measure criterion validity. As it was mentioned through discussion, we used two raters, a physiotherapist and a kinesiologist, who were well experienced on wrist and hand goniometry measurements. Type of experience and also period of practice by the other raters may have affected the force application and the results of the similar studies especially for torque measurements. So, how much the raters of this study may represent other evaluators with different experience and/or period of practice is unknown. Although application of NK and J-Tech electrodigital goniometers may provide more accurate ROM measures for the identified movements of wrist and hand movements; digital goniometers are much expensive compare to traditional simple goniometers. So, the application of digital goniometers may be restricted by the specific cost supportive health centers. A gold standard criterion comparison, like radiography would have allowed us to determine if one device or rater was more accurate than the other. Furthermore, measurements of the both affected and unaffected wrist/hand may have provided an unaffected control for each subject and let to compare reliability in both affected and unaffected joints.

The results of this study may help the clinicians to have a viable accurate alternative method for traditional goniometry. The clinicians may able to have much accurate measurements for joint ROMs that may lead to better assessment of patients' status and progress. The accurate and reliable goniometric measurements can be used to determine impairment ratings and functional progress and provide appropriate decision making.

CONCLUSION

Digital goniometric instruments (NK and J-Tech) demonstrated high reliability coefficients and tight error margins in active wrist ROM and active or passive PIP index flexion in patients with limited wrist and/or hand motion. There was a substantial statistical difference in force application between the raters when performing passive ROM measures for PIP index, but the same rater produced consistent force. However, this difference in force application between raters had a relatively small impact on reliability measurements, since reliability coefficients were high for both instruments. The relationship between individual joint motions obtained by digital goniometric instruments (NK and J-Tech) and patient self-rated pain and function scores (quick DASH and PRWE) suggesting that joint motion impairments contributes to functional disability.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Web site along with the published article.

REFERENCES

[1]   Jones E, Hanly JG, Mooney R, *et al.* Strength and function in the normal and rheumatoid hand. J Rheumatol 1991; 18(9): 1313-8.
        [PMID: 1757930]

[2]   McArthur PA, Milner RH. A prospective randomized comparison of Sutter and Swanson silastic spacers. J Hand Surg 1998; 23(5): 574-7.
        [http://dx.doi.org/10.1016/S0266-7681(98)80004-1] [PMID: 9821596]

[3]   Lefevre-Colau MM, Poiraudeau S, Oberlin C, *et al.* Reliability, validity, and responsiveness of the modified Kapandji index for assessment of functional mobility of the rheumatoid hand. Arch Phys Med Rehabil 2003; 84(7): 1032-8.
        [http://dx.doi.org/10.1016/S0003-9993(03)00128-X] [PMID: 12881830]

[4]   Horger MM. The reliability of goniometric measurements of active and passive wrist motions. Am J Occup Ther 1990; 44(4): 342-8.
        [http://dx.doi.org/10.5014/ajot.44.4.342] [PMID: 2330964]

[5]   Lastayo PC, Wheeler DL. Reliability of passive wrist flexion and extension goniometric measurements: a multicenter study. Phys Ther 1994; 74(2): 162-74.
[PMID: 8290621]

[6]   Ellis B, Bruton A. A study to compare the reliability of composite finger flexion with goniometry for measurement of range of motion in the hand. Clin Rehabil 2002; 16(5): 562-70.
[http://dx.doi.org/10.1191/0269215502cr513oa] [PMID: 12194627]

[7]   Jonsson P, Johnson PW. Comparison of measurement accuracy between two types of wrist goniometer systems. Appl Ergon 2001; 32(6): 599-607.
[http://dx.doi.org/10.1016/S0003-6870(01)00036-9] [PMID: 11703046]

[8]   Armstrong AD, MacDermid JC, Chinchalkar S, Stevens RS, King GJ. Reliability of range-of-motion measurement in the elbow and forearm. J Shoulder Elbow Surg 1998; 7(6): 573-80.
[http://dx.doi.org/10.1016/S1058-2746(98)90003-9] [PMID: 9883416]

[9]   Jonsson P, Johnson PW, Hagberg M. Accuracy and feasibility of using an electrogoniometer for measuring simple thumb movements. Ergonomics 2007; 50(5): 647-59.
[http://dx.doi.org/10.1080/00140130601164490] [PMID: 17454085]

[10]   Breger-Lee D, Voelker ET, Giurintano D, Novick A, Browder L. Reliability of torque range of motion: a preliminary study. J Hand Ther 1993; 6(1): 29-34.
[http://dx.doi.org/10.1016/S0894-1130(12)80178-5] [PMID: 8343872]

[11]   Roberson L, Giurintano DJ. Objective measures of joint stiffness. J Hand Ther 1995; 8(2): 163-6.
[http://dx.doi.org/10.1016/S0894-1130(12)80315-2] [PMID: 7550628]

[12]   Reese NB, Bandy WD. Joint Range of Motion and Muscle Length Testing. Philadelphia: WB Saunders Co 2002.

[13]   Norkin CC, White DJ. Measurement of Joint Motion, A Guide to Goniometry. Philadelphia: FA Davis Co. 2009.

[14]   Marks RG. Analyzing Research Data: The Basis of Biomedical Research Methodology. London, England: Lifetime Learning 1982.

[15]   Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86(2): 420-8.
[http://dx.doi.org/10.1037/0033-2909.86.2.420] [PMID: 18839484]

[16]   Streiner DL, Norman GR. Reliability. In: Streiner DL, Norman GR, Eds. Health Measurement Scales: A practical Guide to their Development and Use. Oxford: Oxford University Press 1995.

[17]   Portney LG, Watkins MP. Foundamentals of Clinical Research, Application to Practice. 3rd ed. New Jersey: Pearson Prentice Hall 2009.

[18]   Fleiss JL. The Design and Analysis of Clinical Experiments. Toronto: John Wiley and Sons 1986.

[19]   Domholdt E. Physical Therapy Research: Principles and Applications. Philadelphia: WB Saunders Co. 1993.

[20]   Beaton DE, Bombardier C, Katz JN, *et al.* Looking for important change / differences in studies of responsiveness. Outcome measures in rheumatology. Minimal clinically important difference. J Rheumatol 2001; 28: 400-5.
[PMID: 11246687]

[21]   Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476): 307-10.
[http://dx.doi.org/10.1016/S0140-6736(86)90837-8] [PMID: 2868172]

[22]   Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998; 26(4): 217-38.
[http://dx.doi.org/10.2165/00007256-199826040-00002] [PMID: 9820922]

[23]   Mason RO, Lind DA, Marchal WG. Statistics: An Introduction. New York: Harcourt Brace Jovanovich Inc. 1983.

[24]   Ludbrook J. Confidence in Altman-Bland plots: a critical review of the method of differences. Clin Exp Pharmacol Physiol 2010; 37(2): 143-9.
[http://dx.doi.org/10.1111/j.1440-1681.2009.05288.x] [PMID: 19719745]

[25]   Bland JM, Altman DG. Comparing two methods of clinical measurement: a personal history. Int J Epidemiol 1995; 24(Suppl. 1): S7-S14.
[http://dx.doi.org/10.1093/ije/24.Supplement_1.S7] [PMID: 7558556]

[26]   Adams BD, Grosland NM, Murphy DM, McCullough M. Impact of impaired wrist motion on hand and upper-extremity performance(1). J Hand Surg Am 2003; 28(6): 898-903.
[http://dx.doi.org/10.1016/S0363-5023(03)00424-6] [PMID: 14642503]

[27]   Kasapinova K, Kamiloski V. Outcome evaluation in patients with distal radius fracture. Contrib, Sect Biol Med Sci 2011; 32(2): 231-46.
[PMID: 22286627]

[28]   MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH. Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. J Hand Surg Am 2000; 25(2): 330-40.
[http://dx.doi.org/10.1053/jhsu.2000.jhsu25a0330] [PMID: 10722826]

[29]   MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. J Orthop Trauma 1998; 12(8): 577-86.
[http://dx.doi.org/10.1097/00005131-199811000-00009] [PMID: 9840793]

[30]   Karnezis IA, Fragkiadakis EG. Association between objective clinical variables and patient-rated disability of the wrist. J Bone Joint Surg Br 2002; 84(7): 967-70.
[http://dx.doi.org/10.1302/0301-620X.84B7.12673] [PMID: 12358387]

[31]   Harris JE, MacDermid JC, Roth J. The International Classification of Functioning as an explanatory model of health after distal radius fracture: a cohort study. Health Qual Life Outcomes 2005; 3: 73.
[http://dx.doi.org/10.1186/1477-7525-3-73] [PMID: 16288664]

[32]   MacDermid JC, Donner A, Richards RS, Roth JH. Patient *versus* injury factors as predictors of pain and disability six months after a distal radius fracture. J Clin Epidemiol 2002; 55(9): 849-54.
[http://dx.doi.org/10.1016/S0895-4356(02)00445-6] [PMID: 12393071]