SCIENTIFIC ARTICLE

Augmented Reality-Based Finger Joint Range of Motion Measurement: Assessment of Reliability and Concurrent Validity

Sasha Gabrielle Létourneau, MD, MSc,* Helen Jin, BSc, MSc,† Ethan Peters, BCS,‡ Ruby Grewal, MD, MSc,§ Douglas Ross, MD, MEd,* Caitlin Symonette, MD, MSc*

Purpose This study aimed to determine the reliability and concurrent validity of finger joint range of motion (ROM) measurement using augmented reality (AR)-based hand tracking in a sample of healthy hands. Additionally, the study aimed to determine which camera view of the hand provided ROM measurements with the highest concurrent validity at each joint.

Methods A web application developed for smart devices using Google's MediaPipe *Hands* framework converted AR-generated hand landmark coordinates from camera feed into ROM angle measurements in real time for all joints. From each of five camera views, we recorded five sets of AR-based flexion and extension measurements at the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints of normal index to small fingers. Test—retest reliability of the five AR-based measurements in each view was evaluated as was concurrent validity of AR-based measurements relative to manual goniometry, considered the reference standard. Given accepted inter-rater reliability of manual goniometry is 10°, we considered AR-based measurements within 10° of goniometry measurements to have "acceptable" concurrent validity.

Results Forty-eight healthy hands (median age 31, 50% left, varying ethnicities) were measured. All joints demonstrated excellent test—retest reliability (intraclass correlation coefficient >0.75) in all views in flexion and ≥2 views in extension. AR-based flexion measurements were within 10° of goniometry in ulnar views of the index MCP, PIP, and DIP; the long MCP and PIP; and the ring PIP and DIP. In extension, multiple views at each joint consistently yielded AR-based measurements within 10° of goniometry.

Conclusions AR-based measurement has high concurrent validity and reliability; however, optimal camera views vary joint to joint. Validation in pathologic hands is required.

Clinical Relevance Given its excellent reliability, AR-based measurement has potential for use in monitoring changes in finger ROM after intervention, either by clinicians in-person or by patients performing remote measurements independently. *(J Hand Surg Am. 2024;■(■):1.e1-e13.*

Key words Augmented reality, machine learning, range of motion measurement, surgical innovation, telemedicine.

Online Exclusive

*From the *Division of Plastic and Reconstructive Surgery, Department of Surgery, Western University, London, Ontario, Canada; †Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada; ‡Western University, London, Ontario, Canada; and §Division of Orthopedic Surgery, Department of Surgery, Western University, London, Ontario, Canada.*

Received for publication December 27, 2023; accepted in revised form October 16, 2024.

Corresponding author: Dr. Sasha Létourneau, 268 Grosvenor St, Room D0-209, London, Ontario, Canada, N6A 4V2; e-mail: sletourneau@qmed.ca.

0363-5023/24/■■-0001
https://doi.org/10.1016/j.jhsa.2024.10.006

THE WIDESPREAD INTEGRATION of technology into health care delivery during the coronavirus disease (COVID)-19 pandemic demonstrated how technology could improve health care delivery, including the care of patients with upper-extremity pathology. Concomitantly, there has been a growing body of literature regarding smartphone-based photography and applications as alternatives to manual goniometry for measuring finger joint range of motion (ROM).[1–6] Furthermore, advances in open-source machine learning (ML) algorithms and augmented reality (AR) frameworks have provided opportunities for development of real-time finger joint tracking tools.[7–9]

We developed a web application (https://digitsrehab.com) for finger joint ROM measurement leveraging MediaPipe Hands, an open-source ML-based cross-platform computer vision model.[8,10] MediaPipe Hands uses palm detector and hand landmark models to extrapolate 21 three-dimensional hand landmark coordinates, each representing a fingertip, interphalangeal (IP) joint, metacarpophalangeal (MCP) joint, thumb carpometacarpal (CMC) joint or wrist coordinate (Fig. 1A).[8] MediaPipe Hands was trained on >10,000 real-world images and >100,000 computer-generated videos with varying backgrounds, gestures, and skin tones.[8] The web application measures ROM in all hand joints simultaneously in real time using vectors between adjacent joint landmarks (Fig. 1B).[7,11] We have previously demonstrated this application's ability to track hands across different lighting, devices, and

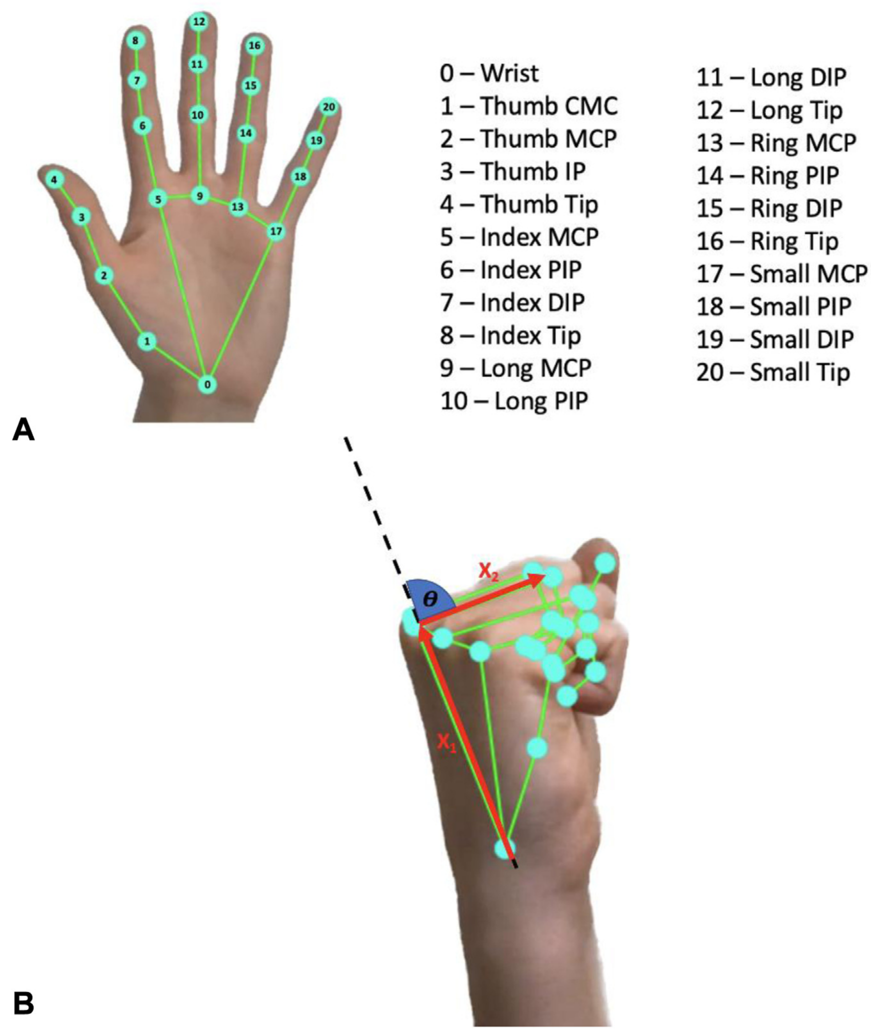

FIGURE 1: A Position and reference locations of the 21 2.5-dimensional hand AR landmarks detected by the MediaPipe Hands model. (Adapted from: https://developers.google.com/mediapipe/solutions/vision/hand_landmarker.). **B** Diagram demonstrating long MCP joint ROM calculation using the wrist and index MCP landmarks for vector X_1 and the index MCP and PIP joint landmarks for vector X_2. The angle at the intersection of X_1 and X_2 is the joint angle (Θ). CMC, carpometacarpal joint; DIP, distal interphalangeal joint; MCP, metacarpophalangeal joint; PIP, proximal interphalangeal joint.

camera resolutions.[9,12] Ultimately, after real-world validation with clinical populations, such a tool could be used, for example, by patients for at-home ROM progress tracking or by providers performing multi-joint ROM measurements in the clinic. However, prior to consideration of real-world trials, AR-based measurements must be determined to be reliable and have acceptable concurrent validity relative to manual goniometry, the reference standard.[13]

Moreover, unlike multi-camera depth-based systems, which require users to purchase expensive hardware and software, this application receives input from only a single RGB camera integrated into most smart devices, which are owned by a majority of Americans, including those of lower socio-economic status.[14–18] The application uses ML pose estimation algorithms to convert two-dimensional input into three-dimensional coordinates.[8] Thus, landmark position based on pose estimation can vary depending on the camera's orientation relative to the hand.[9] Consequently, as part of this validation study, it was also necessary to determine which camera angle provides the most accurate ROM measurements relative to manual goniometry (hereafter "goniometry") at each joint.

This study aimed to assess the reliability and concurrent validity of AR-based finger joint ROM measurement in healthy hands as a first step in testing a potentially clinically relevant AR-based ROM measurement tool. Specifically, we aimed to determine the test–retest reliability of AR-based measurements, the concurrent validity of AR-based measurements from five camera views of the hand relative to manual goniometry, and the camera view providing the highest concurrent validity at each joint.

MATERIALS AND METHODS

Study design

This study assessed reliability and concurrent validity of AR-based finger joint ROM measurement in the index, long, ring, and small fingers from five camera views compared to goniometry. Given the thumb's unique biomechanics, this study focused on the index to small fingers only. For reference, "reliability" refers to the consistency of successive measurements on the same subject under similar conditions. "Concurrent validity" refers to the degree of agreement between a new measurement tool or technique and the reference standard when both measurements are performed at approximately the same time. This study was approved by the University of Western Ontario's Human Ethics Research Board.

Participants

Adults ≥18 years-old had their finger ROM measured in flexion and extension using AR-based and goniometric measurement (Fig. 2). Participants were volunteers consisting of medical trainees, hand clinic staff and their partners. Exclusion criteria included inability to read or follow basic commands in English and any self-reported or visible history of active or prior hand pathology, including conditions such as arthritis, tendon injuries, digit fractures, or Dupuytren disease. Participants consented to video recording of their hands and collection of sociodemographic data intended to ascertain sample diversity, including age, biologic sex, self-identified gender, self-identified ethnicity and hand dominance. Participants were offered $20 CAD compensation to offset parking and travel costs.

Goniometer measurements

A certified hand therapist used a standardized protocol to measure ROM at all joints of the index to small fingers.[19] Each joint was measured once in maximal active flexion and extension using the same goniometer across all participants. By convention, hyperextension values were assigned a negative value.

Video capture of hands for AR-based ROM measurement

A typical set-up for clinical use would involve a single smart device directed at the hand. However, to compare ROM measurements from different camera orientations, we arranged five tablets (iPad Pro 11-inch, ninth generation) in a 10-inch-radius semicircle around the hand on an adjustable-height platform on a desk in a brightly lit room (Fig. 3).[9] Each tablet's front-facing camera was directed centrally to provide a different view of the hand: 10° (ulnar side), 45° (ulnar oblique), 90° (palmar), 135° (radial oblique), and 170° (radial side) (Fig. 3). Splints were used to hold the wrist in neutral (0°) with the forearm perpendicular to the tablet platform. Immediately following goniometry measurement, videos were taken from the five camera views simultaneously using voice-activated recording on the tablets while participants moved their hands into maximal active flexion (tightest composite fist) and extension (Fig. 2). Participants held maximal flexion and extension poses for 3 seconds each with the thumb palmarly abducted and in view of all cameras. A demonstration was provided, and participants were prompted throughout recording.

After capture, videos were edited into 10-second clips of the hand moving into flexion or extension. Each clip showed 7 seconds of the hand moving into

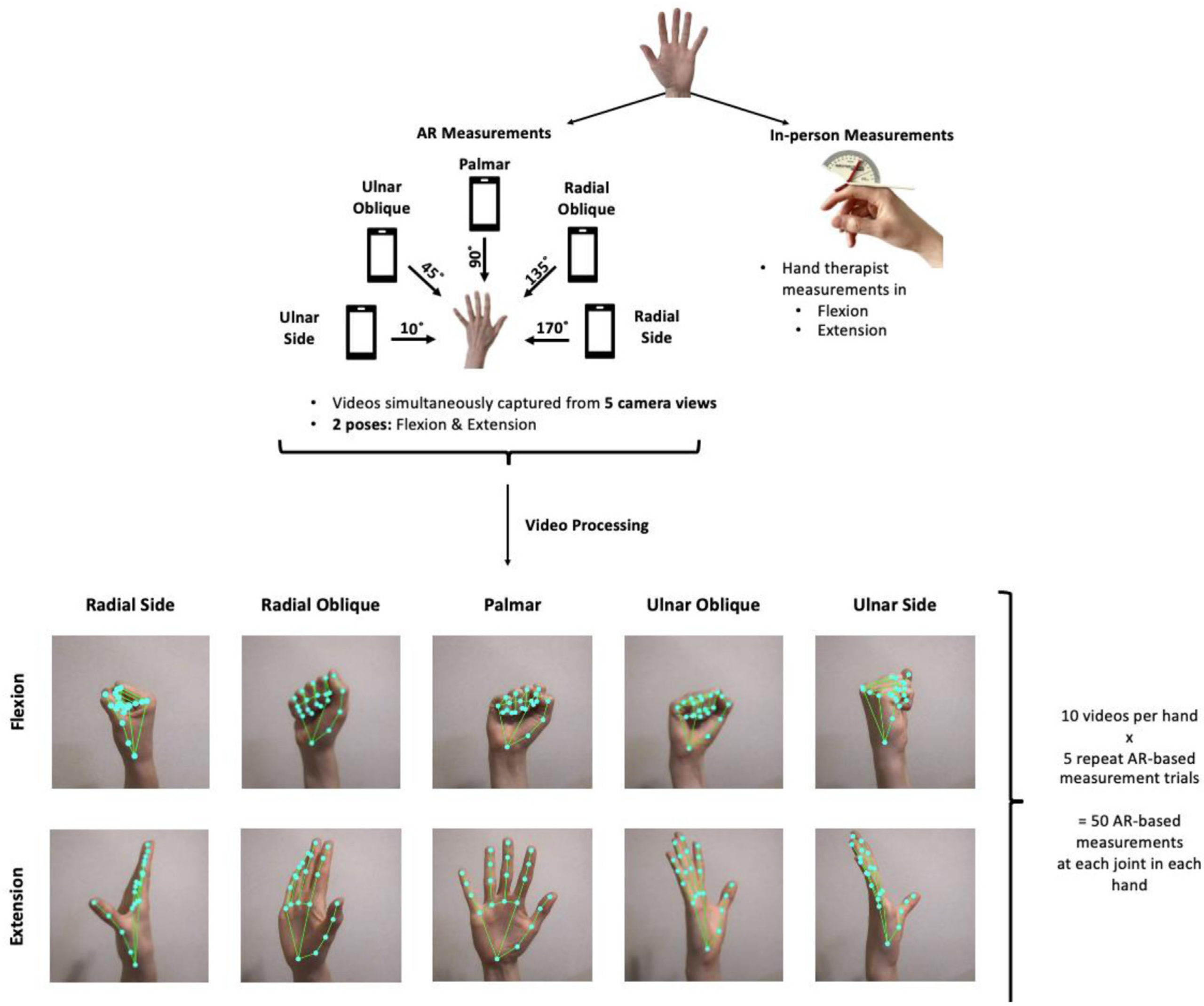

FIGURE 2: Flow diagram demonstrating how in-person and AR-based ROM measurements were acquired.

maximal flexion or extension in the final frame, with a 3-second freeze-frame of the final frame position. Other hands in the background of the video were covered using shape animations in Microsoft PowerPoint (Version 16.83) and final video versions were rerecorded using screen recording on a MacBook Pro.

AR-based ROM measurement

The current functionality of the web application allows real-time ROM angle calculation using algorithms described in our previous studies.[9,12] We also enabled videos to be uploaded to the web application as input for the ROM measurement framework as an alternative to real-time camera feed. We copied each flexion and extension video five times and uploaded these to the AR-based web application on a single computer with stable internet connection. This process allowed capture of five repeated "trials" of each participant's index to small finger joint ROM from each of the five camera views (Fig. 2). Given that the application tracks joints throughout every frame of the video, the average number of ROM measurements performed per video at each joint was 396 (standard deviation 70) in extension and 386 (standard deviation 67) in flexion. The maximal AR-based ROM angle at each joint was extracted from the flexion data sets. Although our application can track joints in hyperextension, it is currently unable to assign negative values to hyperextended ROM measurements, as is the convention with goniometry measurement. Thus, attempting to extract the minimum angle measurement would have yielded the measurement closest to 0° rather than maximal hyperextension, falsely underestimating agreement with goniometry. Consequently, we extracted the last AR-based ROM measurement at each joint in the

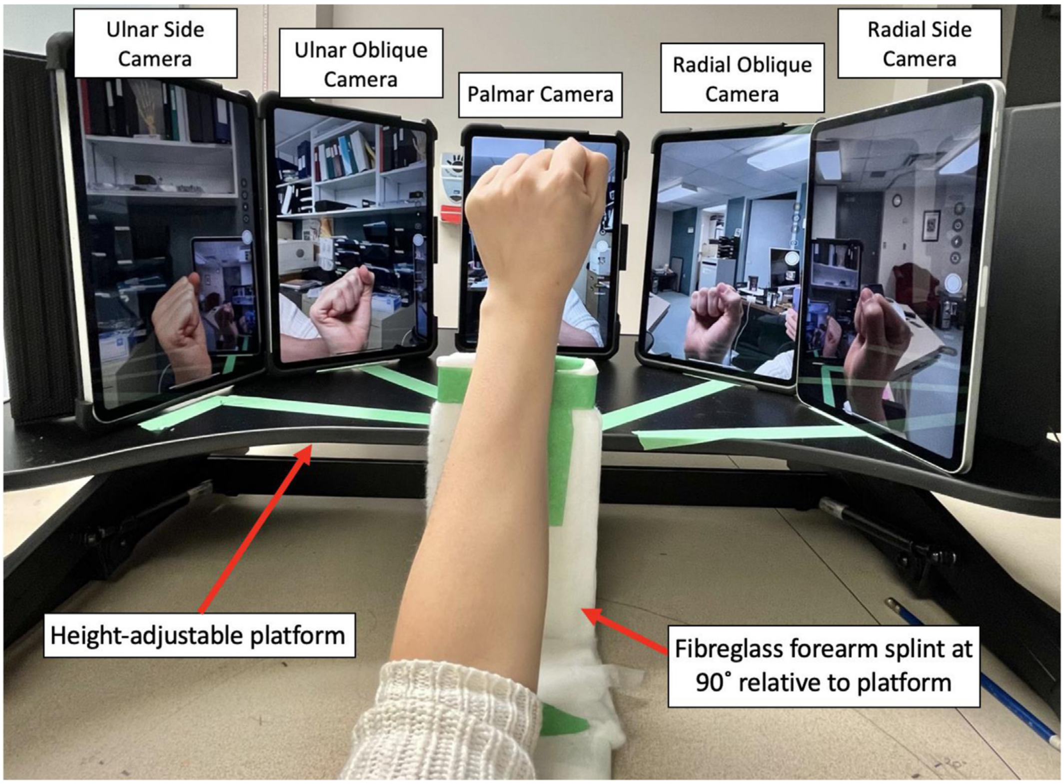

FIGURE 3: Tablet set-up for video capture of the hand from five camera views. Note: The posterior forearm splint is not shown in this photo.

extension data sets for comparison to the absolute in-person measurement.

Data analysis

An *a priori* sample size calculation for a paired *t*-test was performed using alpha of 0.05 and 80% power with a mean difference of 5°, the intrarater reliability of goniometry, and a standard deviation of 10°. This yielded a sample of 34 hands; however, we recruited a larger sample anticipating potential technologic issues.

We analyzed the aforementioned sociodemographic variables using descriptive analysis and unweighted frequencies. Age was not normally distributed (Shapiro−Wilks test, $P < .05$) and, therefore, summarized using median and interquartile range (IQR).

Reliability

Intraclass correlation coefficients (ICC) were used to assess test−retest reliability of the five repeated AR-based measurements at each joint in each camera view. Single measures ICC values using a two-way mixed effects model with an absolute agreement definition (ICC 3,1) were reported as is the convention for test−retest reliability.[20−22] Cicchetti and Sparrow[23] define ICC under 0.40 as "poor", 0.4−0.59 as "fair", 0.60−0.74 as "good", and 0.75−1.00 as "excellent" reliability.[4] We calculated standard error of the measurement (SEM) of the AR-based measurement tool at each joint in each camera view to provide an estimate of the tool's measurement error. SEM was calculated using the square root of the mean square error from the two-way analysis of variance used in the ICC calculation.[21] Minimal detectable change (MDC_{90}) was calculated with a confidence of 90% using the SEM value.[22,24] The MDC_{90} estimates the amount of change in ROM values required to be 90% confident a change in joint ROM has actually occurred.

Concurrent validity

The mean of the five AR-based measurements from each camera view were compared to the goniometry measurement at each joint using Wilcoxon signed-rank tests. Differences between measurement techniques were reported using median and IQR (Shapiro−Wilks test, $P < .05$). In extension, absolute goniometry measurements were used, as the application cannot automatically assign negative values in hyperextension.[25,26] Means of the AR-based and goniometry measurements are plotted against their

TABLE 1. Participant Sociodemographic Data

Sociodemographic Data	Value
Participants (n)	24
Age (y, median [range])	31.5 (23−54)
Woman-identifying (n [%])	13 (54.2)
Self-identified ethnic background (n [%])	
Arabic	1 (4.2)
Caucasian (European descent)	9 (37.5)
Indigenous, First Nations (status or nonstatus), Métis, Inuk (Inuit)	1 (4.2)
Latin American	2 (8.3)
South Asian (eg, East Indian, Pakistani, Sri Lankan, etc)	2 (8.3)
South-East Asian (Chinese, Korean, Japanese, Filipino, Vietnamese, Cambodian, Laotian, Thai, etc)	7 (29.2)
West Asian (eg, Iranian, Afghan, etc)	2 (8.3)

differences to produce Bland-Altman plots with 95% limits of agreement (LOAs). The proportion of AR-based measurements falling within 10° of goniometry measurements at each joint and in each camera view is reported.[27] The 10° threshold represents the inter-rater reliability of goniometry, the clinically important difference threshold used in other studies.[28,29]

RESULTS

Twenty-four individuals without hand pathology had both hands measured, resulting in a sample of 48 hands. The radial oblique camera failed to record one participant's hands bilaterally. The ulnar side camera failed to record another participant's left hand, resulting in 48 flexion and extension videos each from the palmar, radial side and ulnar oblique views, 47 videos from the ulnar side view and 46 from the radial oblique view. Participant sociodemographic data are presented in Table 1.

AR-based measurement test−retest reliability

Table 2 summarizes the camera view yielding the highest test−retest reliability, meaning the highest ICC estimates and lowest SEM and MDC_{90} values at each joint (results in all camera views are available in Supplementary Table 1, available online on the *Journal*'s website at www.jhandsurg.org). In flexion, ICC estimates were consistently excellent (ICC $\geq$ 0.75) in all joints across all views with the ulnar oblique view most often yielding SEM and MDC_{90} values under 5°, the generally accepted intrarater reliability of goniometry.[19] In extension, ICC values varied more widely, with all joints demonstrating two or more views with excellent reliability. In contrast to flexion measurements, SEM and MDC_{90} in extension were most consistently less than 5° in the palmar and radial oblique views.

Concurrent validity relative to goniometry and optimal camera view

Mean AR-based and single goniometry measurements were compared at each joint in each camera view in flexion and extension. In flexion, AR-based and goniometry measurements were within 10° in all index joints in the ulnar side view (Table 3). Additionally, in the long finger, the ulnar oblique view of the MCP and DIP joints and the ulnar side view of the PIP joint were most accurate in flexion (Table 3). Agreement in the ulnar two fingers in flexion was more limited. Specifically, only the ring finger PIP and DIP joints in the ulnar side and ulnar oblique views, respectively, demonstrated median differences less than 10° (Table 3). In extension, multiple camera views at each joint across all fingers yielded a median difference between AR-based and goniometry measurements $\leq$10° (Supplementary Table 2 summarizes median differences in all camera views, available online on the *Journal*'s website at www.jhandsurg.org). The camera view with the smallest median difference between techniques (Table 3) was often the view where the highest proportion of AR-based measurements fell within 10° of goniometry measurements (Table 4, see Supplementary Table 3 for results across all camera views, available online on the *Journal*'s website at www.jhandsurg.org). In flexion, the proportion of AR measurements within 10° of in-person measurements ranged from 37.0% to 72.3% in the best camera view compared to 54.2% to 89.6% in extension. In extension, again, multiple camera views yielded high proportions of AR-based measurements within 10° of in-person measurements. Figure 4 demonstrates Bland-Altman plots for the index with other fingers available in Supplementary Figure S1 (available online on the *Journal*'s website at www.jhandsurg.org).

DISCUSSION

This study assessed test−retest reliability and concurrent validity of AR-based finger joint ROM measurement relative to goniometry in a cohort without hand pathology. When presented with identical videos of the same hand, AR-based measurement demonstrated excellent reliability across multiple camera views. Further, concurrent validity depended

TABLE 2. **Summary of the Most Reliable Camera Views at Each Joint in Each Finger as Determined by Relative and Absolute Measures of Test−retest Reliability of Five Repeated Trials of AR-based Measurements**

		Flexion					Extension				
Finger	Joint	Most Reliable Camera View of the Hand	ICC (3,1) for Single Measures	95% CI	SEM	MCD_{90}	Most Reliable Camera View of the Hand	ICC (3,1) for Single Measures	95% CI	SEM	MCD_{90}
Index	MCP	Ulnar oblique	1.00	1.00−1.00	0.5	1.1	Palmar	0.99	0.99−1.00	0.7	1.5
	PIP	Ulnar oblique	0.97	0.95−0.98	2.0	4.6	Radial side	0.83	0.76−0.89	0.7	1.5
	DIP	Radial oblique	0.93	0.90−0.96	8.8	20.4	Radial oblique	0.99	0.98−0.99	0.5	1.1
Long	MCP	Ulnar oblique	1.00	1.00−1.00	0.5	1.1	Radial oblique	0.99	0.98−0.99	0.7	1.6
	PIP	Ulnar oblique	0.92	0.88−0.95	2.4	5.5	Palmar	0.99	0.99−1.00	0.9	2.0
	DIP	Palmar	0.98	0.96−0.98	3.2	7.5	Radial oblique	0.96	0.94−0.98	0.8	1.7
Ring	MCP	Ulnar oblique	1.00	1.00−1.00	0.5	1.3	Radial oblique	0.98	0.97−0.99	1.0	2.3
	PIP	Ulnar oblique	0.98	0.97−0.99	1.4	3.3	Radial oblique	0.96	0.94−0.98	1.5	3.4
	DIP	Ulnar oblique	0.98	0.98−0.99	1.2	2.7	Radial oblique	0.94	0.90−0.96	1.0	2.3
Small	MCP	Ulnar oblique	1.00	1.00−1.00	0.7	1.5	Palmar	0.99	0.98−0.99	1.0	2.8
	PIP	Ulnar oblique	0.99	0.98−0.99	1.6	3.7	Palmar	0.98	0.97−0.99	0.7	1.6
	DIP	Palmar	0.99	0.98−0.99	2.4	5.7	Palmar	0.98	0.97−0.99	0.5	1.1

TABLE 3. Summary of the Camera View by Joint Yielding the Lowest Absolute Difference Between AR-based and In-person Measurements

		Flexion				Extension			
Finger	Joint	Camera View of the Hand	Median of Absolute Differences	IQR (Q1−Q3)	Wilcoxon Signed-rank Test (*P*)	Camera View of the Hand	Median of Absolute Differences	IQR (Q1−Q3)	Wilcoxon Signed-rank Test (*P*)
	MCP	Ulnar side	7.9	4.4−14.3	<.001	Radial oblique	8.5	4.0−15.9	.002
	PIP	Ulnar side	6.6	3.6−13.8	.15	Ulnar side	4.1	1.8−9.2	.003
	DIP	Ulnar side	9.4	3.7−14.7	.933	Radial side	3.5	1.9−7.3	.129
Long	MCP	Ulnar oblique	6.4	2.1−13.8	.255	Ulnar oblique	6.8	3.5−13.3	.303
	PIP	Ulnar side	6.8	3.1−9.7	.015	Radial side	6.0	2.7−9.2	.372
	DIP	Ulnar oblique	12.4	6.6−17.9	.001	Palmar	3.2	1.8−7.0	.036
Ring	MCP	Palmar	10.2	5.5−18.2	.002	Ulnar side	9.3	4.6−18.8	.079
	PIP	Ulnar side	6.8	3.0−13.1	.003	Radial side	5.9	3.1−17.6	<.001
	DIP	Ulnar oblique	9.0	4.4−15.3	.025	Palmar	3.3	1.7−6.5	.454
Small	MCP	Palmar	15.6	9.2−27.2	.689	Ulnar oblique	8.3	6.3−17.7	<.001
	PIP	Ulnar side	10.2	3.8−19.5	<.001	Ulnar oblique	5.9	2.6−11.5	<.001
	DIP	Ulnar oblique	16.8	8.6−29.2	<.001	Ulnar side	4.3	2.0−7.1	.368

TABLE 4. Summary of the Camera View by Joint Yielding the Highest Concurrent Validity as Determined by the Bland-Altman Analysis of AR-based Measurements Relative to In-person Measurements

		Flexion					Extension				
Finger	Joint	Camera View of the Hand	Mean of Differences	Lower LOA	Upper LOA	Proportion of AR Measurements Within 10° of In-person	Camera View of the Hand	Mean of Differences	Lower LOA	Upper LOA	Proportion of AR Measurements Within 10° of In-person
Index	MCP	Ulnar side	6.8	-14.1	27.7	59.6	Radial oblique	-6.0	-29.5	17.5	56.5
	PIP	Ulnar side	-3.6	-26.7	19.4	63.0	Radial side	2.3	-15.6	12.1	85.4
	DIP	Ulnar oblique	0.4	-13.2	36.8	60.4	Ulnar side	-0.3	-12.6	12.0	89.4
Long	MCP	Ulnar oblique	1.8	-20.9	24.5	64.6	Ulnar oblique	-1.4	-23.7	20.9	66.7
	PIP	Ulnar side	4.4	-19.7	28.6	72.3	Radial side	2.0	-36.0	40.0	79.2
	DIP	Ulnar oblique	8.8	-24.9	42.4	37.5	Ulnar oblique	-0.7	-14.2	12.8	87.5
Ring	MCP	Palmar	-7.3	-35.6	20.9	47.9	Ulnar side	-2.9	-34.3	28.6	55.3
	PIP	Ulnar side	7.7	0.2	39.5	66.0	Ulnar oblique	-0.2	-19.8	19.5	68.8
	DIP	Ulnar oblique	3.7	-22.8	30.2	58.3	Palmar	-1.4	-14.3	11.5	89.6
Small	MCP	Radial oblique	-8.5	-42.7	25.7	39.1	Ulnar oblique	-7.4	-30.0	15.3	54.2
	PIP	Ulnar side	9.9	-24.9	44.8	48.9	Palmar	0.3	-16.8	17.3	83.3
	DIP	Radial oblique	24.8	-28.2	77.8	37.0	Ulnar side	7.2	-16.1	12.8	87.0

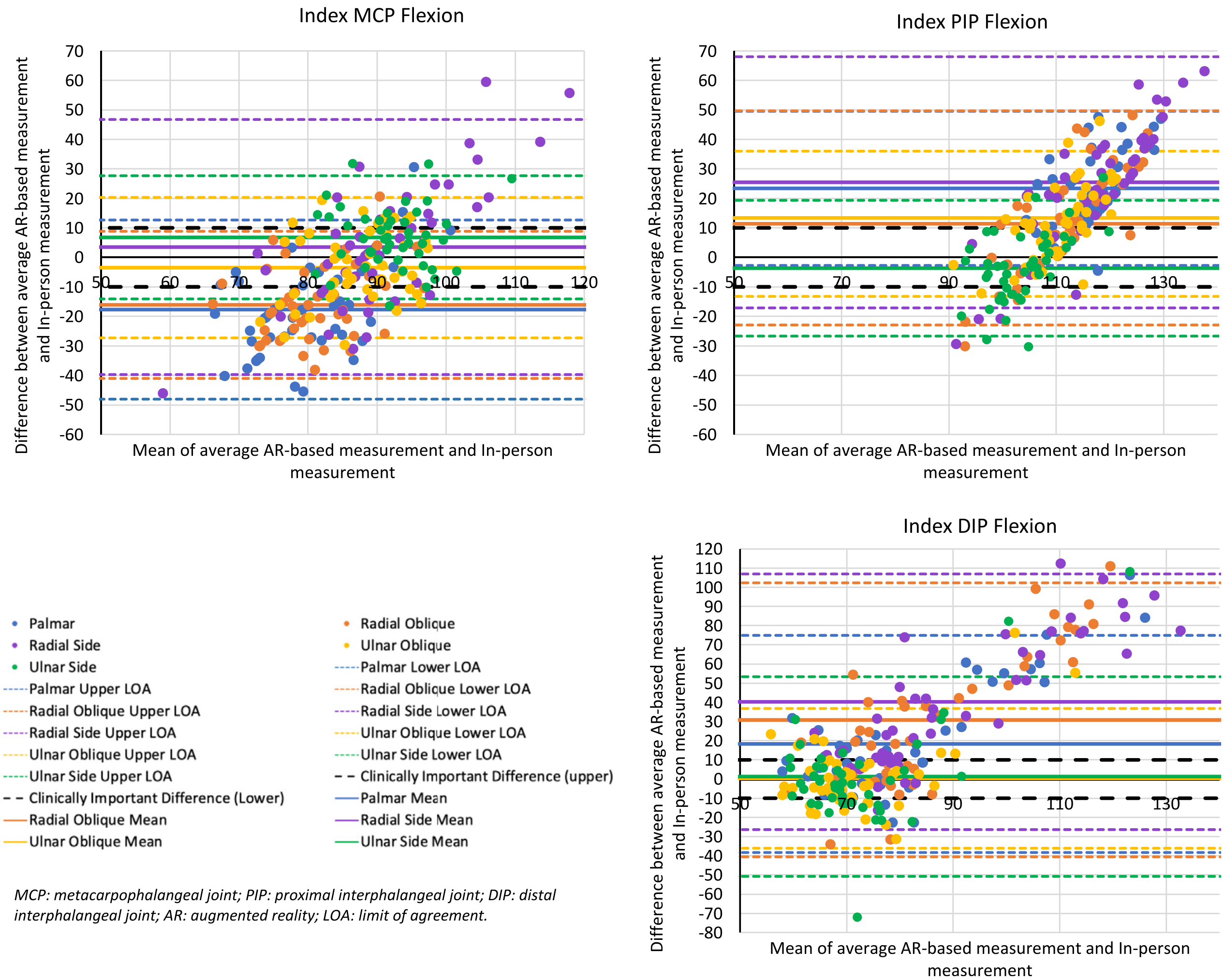

FIGURE 4: Bland-Altman plots comparing average augmented reality-based measurements with goniometry measurements in flexion and in extension in the index finger.

on camera view. Both flexion measurements and extension measurements were often within the accepted inter-rater reliability of goniometry when captured from ulnar side or ulnar oblique views; however, extension measurements, in particular, showed high agreement with goniometry across most camera views.

Although not a feasibility trial for the end-user application, this validation study provides insight into the potential clinical use of AR-based measurement tools by informing optimal hand positioning relative to the camera when measuring ROM on a single smart device. Ultimately, algorithms could be designed to initiate ROM capture only when the hand is positioned correctly in view of the camera. Furthermore, as a feature of a telemedicine-based rehabilitation application, patients could use AR-based measurement to track ROM progression during home therapy and update their providers through the application platform. Indeed, studies across various fields of medicine have linked at-home progress tracking with patient motivation and therapy adherence.[30–37] Such a tool could also flag patients requiring intervention, such as those lagging in their recovery or developing worsening joint contractures. The excellent reliability of AR-based measurement has the potential to facilitate accurate trending of ROM progression. Furthermore, while remote patient-generated ROM trending could possibly reduce the frequency of in-person visits required for ROM measurement, AR-based measurement also offers a potentially efficient alternative to manual goniometry during in-person clinics given it measures all finger joints simultaneously. However, assessment of efficiency was outside the scope of this study.

Although reliability is an important parameter to study when developing new technology to be used

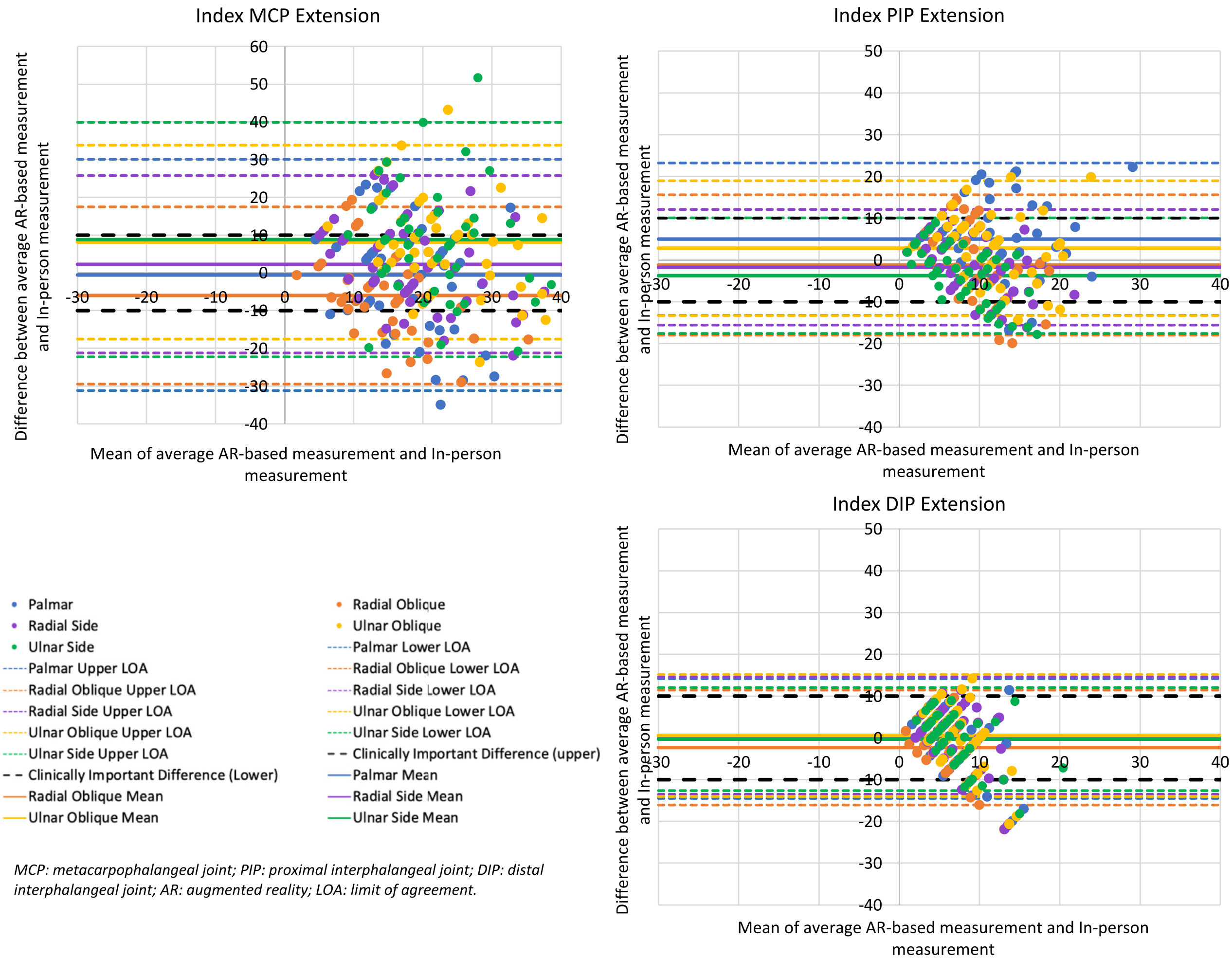

FIGURE 4: (*continued*).

for clinical measurement, it was not evaluated in the few other studies comparing AR-based finger ROM measurement to goniometry.[38,39] Our reliability results were comparable to an accelerometer-based smartphone application (EHMROM; ICC: 0.997), one of the few technology studies that has reported test–retest reliability.[4–6] Our reliability results were also similar to manual goniometry.[29] Lewis et al[40] report comparable or worse intrarater reliability ICC values of 0.64–0.99 for joints measured with a goniometer in active flexion. Measurements of ROM at all joints in our study demonstrated a SEM under 5° in at least one camera view, which compares to previously reported intrarater SEM values of 2° to 4° for goniometry.[41] Finally, our MDC_{90} results showed greater reliability than MDC_{95} values previously reported for goniometry which varied from 12° to 24° across joints.[42] Again, the excellent reliability of AR-based measurement highlights this tool's potential clinical utility for measuring ROM change over time, although validation in pathologic hands is required.

Ulnar views generally provided higher concurrent validity in flexion, whereas optimal views in extension varied by finger. Gu et al[38] also used AR-based measurement to measure ROM from images in a palmar view. They found greater accuracy in extension than in flexion, which is congruent with our findings, as well as higher concurrent validity in PIP and DIP than in MCP joints.[38] We observed that rotating the hand from palmar to ulnar views shifts the MCP landmarks from a more volar to a more dorsal position over the MCP joint. The latter landmark position better approximates a dorsally placed goniometer. Although Gu et al[38] seemingly adjust the MCP joint angle calculation to account for this, our approach yielded higher concurrent validity at all MCP joints, particularly in the index and long fingers. Additionally, ulnar side views likely align the finger and wrist landmarks in similar depth planes, potentially reducing error introduced through pose estimation of the landmark's depth coordinate, which, conversely, is required when joints are measured from the palmar view.

Additionally, although we anticipated higher concurrent validity for ulnar digits in the ulnar views given the camera's unimpeded view of them, the radial digits had higher concurrent validity. A prior study using MediaPipe Hands to recognize abnormal hand postures in nerve injury also noted reduced accuracy in tracking the ring and small fingers.[39] We postulate the framework's single wrist landmark fails to account for the position and increased mobility of the ulnar-sided CMC joints, skewing the wrist-MCP vector used in the MCP joint angle calculation. Further, although not described in the article by Zhang et al,[8] we also hypothesize the radial fingers may have been prioritized when training the MediaPipe Hands' hand-tracking framework, as these fingers are more typically used for gaming gestures like pointing and pinching. This consideration highlights the importance of clinician involvement when designing clinically relevant tools, such as the application proposed in this study.

Our sample size was comparable to similar studies, although our cohort was more diverse in physical characteristics, such as sex, age, and ethnicity, which could potentially affect computer vision.[12,28,38,39] We acknowledge however, the incomplete representation of all ethnicities and that ethnicity is an imperfect proxy for skin tone given variability within ethnic groups. The framework upon which our application was built is trained on >100,000 video recordings of hands with different skin tones[8]; however, future studies should use validated scales, such as the Fitzpatrick or Monk scales, to assess skin tone. Moreover, although beyond the study's scope, feasibility trials of an end-user application should also include metrics of socio-economic diversity, such as income and educational level, to ascertain accessibility of applications such as the one proposed in this study. Another notable strength of this study is that all goniometry measurements were performed by the same therapist using the same goniometer, minimizing confounding errors related to inter-rater or inter-instrument variability.

In terms of limitations, because we intentionally captured measurements from identical videos to remove human error from our test−retest analysis, our results may overestimate the "real-world" reliability of the application, particularly if it were used remotely by patients without a clinician present. Furthermore, this application's inability to label hyperextended joints as negative values necessitated comparison of AR-based measurements to absolute goniometry measurements in extension. Although unlikely, instances where hyperextension was not detected by AR would result in underestimation of the difference between AR-based and goniometry measurements. Future directions include software refinement to enable distinction of hyperextension and validation of AR-based measurement in the thumb and in pathological hands. Feasibility trials of an end-user application with patients and clinicians using the application independently are required to ascertain and optimize application usability.

CONFLICTS OF INTEREST

No benefits in any form have been received or will be received related directly to this article.

ACKNOWLEDGMENTS

The authors would like to acknowledge the efforts of Luann Macdonald, hand therapist, who performed the in-person goniometry measurements for all participants in this study. Additionally, Dr Joy MacDermid, Jacob Davidson, and Dr Claire Parent are acknowledged for their assistance throughout the study. Furthermore, we are grateful to the participants whose participation made this study possible. This research was funded by the New Frontiers in Research Fund (NFRFE-2021-00671) as well the Western University Department of Surgery Internal Research Fund and the Western University Department of Surgery Resident Research Grant. The funding sources had no direct involvement in the conduct of the research, preparation of the article or in the decision to submit the article for publication.

REFERENCES

1. Keogh JWL, Cox A, Anderson S, et al. Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: a systematic review. *PLoS One*. 2019;14(5):e0215806.
2. Longoni L, Brunati R, Sale P, Casale R, Ronconi G, Ferriero G. Smartphone applications validated for joint angle measurement: A systematic review. *Int J Rehabil Res*. 2019;42(1):11−19.
3. Milani P, Coccetta CA, Rabini A, Sciarra T, Massazza G, Ferriero G. Mobile smartphone applications for body position measurement in rehabilitation: a review of goniometric tools. *PM R*. 2014;6(11):1038−1043.
4. Zhao JZ, Blazar PE, Mora AN, Earp BE. Range of motion measurements of the fingers via smartphone photography. *Hand (N Y)*. 2020;15(5):679−685.
5. Miyake K, Mori H, Matsuma S, et al. A new method measurement for finger range of motion using a smartphone. *J Plast Surg Hand Surg*. 2020;54(4):207−214.
6. Lee HH, St. Louis K, Fowler JR. Accuracy and reliability of visual inspection and smartphone applications for measuring finger range of motion. *Orthopedics*. 2018;41(2):e217−e221.
7. Zhang Y. Applications of Google MediaPipe Pose Estimation Using a Single Camera. The California State Univeristy; 2016. Dissertation. Accessed October 8, 2022. https://scholarworks.calstate.edu/concern/projects/2z10ww67b
8. Zhang F, Bazarevsky V, Vakunov A, et al. Mediapipe hands: on-device real-time hand tracking. *ArXiv Prepr ArXiv*. 2020;2006(10214).

9. Dong H, Ho E, Shin H, et al. 'DIGITS' app - smartphone augmented reality for hand telerehabilitation. *Comput Methods Biomech Biomed Eng Imaging Vis*. 2022;10(4):375−382.
10. Lugaresi C, Tang J, Nash H, et al. Mediapipe: A framework for building perception pipelines. *ArXiv*. 2019;1906(08172).
11. Renotte N. YouTube. AI hand pose estimation with mediapipe | detect left and right hand + calculate angles. Accessed September 25, 2022. https://www.youtube.com/watch?v=EgjwKM3KzGU
12. Dong H, Shin H, Ho E, et al. Next-generation remote hand assessments: cross-platform DIGITS web application. *J Hand Surg Glob Online*. 2023;5(3):294−299.
13. Groth GN, Ehretsman RL. Goniometry of the proximal and distal interphalangeal joints, part I: A survey of instrumentation and placement preferences. *J Hand Ther*. 2001;14(1):18−22.
14. Nizamis K, Rijken NHM, Mendes A, Janssen MMHP, Bergsma A, Koopman BFJM. A novel setup and protocol to measure the range of motion of the wrist and the hand. *Sensors (Basel)*. 2018;18(10):1−14.
15. Metcalf CD, Robinson R, Malpass AJ, et al. Markerless motion capture and measurement of hand kinematics: validation and application to home-based upper limb rehabilitation. *IEEE Trans Biomed Eng*. 2013;60(8):2184−2192.
16. HoloKit Docs. Advanced Use of Hand Tracking. Accessed July 19, 2023. https://docs.holokit.io/for-developers/holokit-unity-sdk/hand-tracking
17. Shandhi MMH, Singh K, Janson N, et al. Assessment of ownership of smart devices and the acceptability of digital health data sharing. *NPJ Digit Med*. 2024;7(1):44.
18. Sidoti O, Pew Research Center. Mobile Fact Sheet, 2024. Accessed April 13, 2024. https://www.pewresearch.org/internet/fact-sheet/mobile/?tabId=tab-64e32376-5a21-4b1d-8f8b-5f92406db984
19. Adams L, Greene L, Topoozian E. Range of motion. In: Valdes K, MacDermid J, Solomon G, eds. *Clinical assessment recommendations*. 2nd ed. American Society of Hand Therapists; 1992:55−69.
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155−163.
21. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. *J Strength Cond Res*. 2005;19(1):231−240.
22. Portney LG, Watkins MP. Chapter 26: Statistical Measures of Reliability. In: *Foundations of Clinical Research: Applications to Practice*. 2nd ed. Prentice Hall; 2000.
23. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981;86(2):127−137.
24. Chen S, Wolf SL, Zhang Q, Thompson PA, Winstein CJ. Minimal detectable change of the actual amount of use test and the motor activity log: the EXCITE trial. *Neurorehabil Neural Repair*. 2012;26(5):507−514.
25. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology*. 2010;73(9):1167−1179.
26. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48(6):661−671.
27. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8(2):135−160.
28. Chen J, Xian Zhang A, Jia Qian S, Jing Wang Y. Measurement of finger joint motion after flexor tendon repair: smartphone photography compared with traditional goniometry. *J Hand Surg Eur Vol*. 2021;46(8):825−829.
29. van Kooij YE, Fink A, Nijhuis-van der Sanden MW, Speksnijder CM. The reliability and measurement error of protractor-based goniometry of the fingers: A systematic review. *J Hand Ther*. 2017;30(4):457−467.
30. Argent R, Daly A, Caulfield B. Patient involvement with home-based exercise programs: can connected health interventions influence adherence? *JMIR MHealth UHealth*. 2018;6(3):e47.
31. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: A systematic review of the literature. *J Am Diet Assoc*. 2011;111(1):92−102.
32. Evans J, Papadopoulos A, Silvers CT, et al. Remote health monitoring for older adults and those with heart failure: adherence and system usability. *Telemed J E Health*. 2016;22(6):480−488.
33. Günther G, Junker R, Sträter R, et al. Symptomatic ischemic stroke in full-term neonates : role of acquired and genetic prothrombotic risk factors. *Stroke*. 2000;31:2437−2441.
34. Kim JY, Wineinger NE, Steinhubl SR. The Influence of wireless self-monitoring program on the relationship between patient activation and health behaviors, medication adherence, and blood pressure levels in hypertensive patients: a substudy of a randomized controlled trial. *J Med Internet Res*. 2016;18(6):e116.
35. Lu Z, Signer T, Sylvester R, Gonzenbach R, Von Wyl V, Haag C. Implementation of remote activity sensing to support a rehabilitation aftercare program: observational mixed methods study with patients and health care professionals. *JMIR MHealth UHealth*. 2023;11:e50729.
36. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. *J Am Geriatr Soc*. 2003;51(3):387−392.
37. Trefond J, Hermet L, Lambert C, et al. Home blood pressure monitoring and adherence in patients with hypertension on primary prevention treatment: a survey of 1026 patients in general medicine in the Auvergne region. *BMC Prim Care*. 2022;23(1):131.
38. Gu F, Fan J, Wang Z, Liu X, Yang J, Zhu Q. Automatic range of motion measurement via smartphone images for telemedicine examination of the hand. *Sci Prog*. 2023;106(1):368504231152740.
39. Gu F, Fan J, Cai C, et al. Automatic detection of abnormal hand gestures in patients with radial, ulnar, or median nerve injury using hand pose estimation. *Front Neurol*. 2022;13:1052505.
40. Lewis E, Fors L, Tharion WJ. Interrater and intrarater reliability of finger goniometric measurements. *Am J Occup Ther*. 2010;64(4):555−561.
41. Macionis V. Reliability of the standard goniometry and diagrammatic recording of finger joint angles: a comparative study with healthy subjects and non-professional raters. *BMC Musculoskelet Disord*. 2013;14:17.
42. Reissner L, Fischer G, List R, Taylor WR, Giovanoli P, Calcagni M. Minimal detectable difference of the finger and wrist range of motion: comparison of goniometry and 3D motion analysis. *J Orthop Surg Res*. 2019;14(1):173.