JRSR 4 (2016) 77-84

Journal of Rehabilitation Sciences and Research

Journal Home Page: jrsr.sums.ac.ir

Original Article

The Reliability of Smartphone and Goniometric Measurements of Hip Range of Motion

Elizabeth S Norris[1]*, Emily Wright[2], Shannon Sims[3], Megan Fuller[4], Kurt Neelly[1]

[1]Associate Professor, Department of Physical Therapy, Western Kentucky University, United States
[2]Physical Therapist, Total Fitness Connection, United States
[3]Physical Therapist, The Medical Center Home Care Program, United States
[4]Physical Therapist, Signature Healthcare of Bowling Green, United States

ARTICLE INFO

Article History:
Received: 18/9/2016
Revised: 24/10/2016
Accepted: 30/10/2016

Keywords:
Hip
Goniometer
Smartphone
Range of motion
Reliability

ABSTRACT

Background: Range of motion is an essential component of the hip examination. Handling issues with the goniometer often create challenges when measuring hip passive range of motion (PROM). Recent generations of smartphones have emerged as an alternative instrument for the measurement of joint ROM. The purpose of this study was to investigate the intra-rater, inter-rater and inter-instrument reliability of smartphone and goniometric hip PROM.

Methods: Two investigators measured hip PROM to a designated end position on 30 asymptomatic participants in a blinded within study design using two measurement methods, smartphone and goniometer. Relative reliability of smartphone and goniometric measurements of hip PROM was assessed using intraclass correlation coefficients (ICC). Absolute reliability of both measurement methods was assessed using paired t-tests, standard errors of measurement (SEM), and 95% limits of agreement (LOA).

Results: Relative reliability ICCs ranged from 0.47-0.99 (intra-rater), 0.05-0.99 (inter-rater) and 0.25 -0.97 (inter-instrument). Inter-rater differences of smartphone hip measurements were non-significant, however, significant differences were found for all inter-rater goniometric hip measurements ($P<0.02$). The comparison of the smartphone to goniometric measurements showed bias was present in 7 out to 12 hip measurements ($P<0.04$). SEM ranged from 1° to 3° (intra-rater, inter-rater, and inter-instrument). LOA ranged from -6.8° - 5.1° (inter-rater) and -8.9° - 13.8° (inter-instrument).

Conclusion: These findings support intra-rater reliability of both instruments when measuring hip PROM. Inter-rater reliability, however, was supported only for the smartphone. Due to systematic bias of inter-instrument measurements performed by one rater, caution should be used if the instruments are to be used interchangeably in order to quantify within session hip PROM.

Introduction

Measurement of hip range of motion (ROM) is an important component in the management of lower quarter dysfunction. Hip ROM is used to quantify pathology related impairments of mobility, monitor disease progression, determine the effectiveness of rehabilitation interventions, and as a variable in clinical predication rules [1–3]. The instruments most commonly used in the clinic to measure hip ROM are the goniometer and the

*Corresponding author: Elizabeth S Norris, Western Kentucky University, bowling green, United States.
Tel: +270-745-3188
E-mail: elizabeth.norris@wku.edu

inclinometer [2,4,5]. While both instruments are portable and easy to use, they each have associated limitations. Goniometric hip measurements generally have lower inter-rater reliability (ICC=0.22-0.87)[2,6]. In addition, there are challenges in the ability to correctly align the goniometer while handling the lower extremity (LE) while providing stabilization to other areas during hip PROM measurements [2,7]. While use of the inclinometer minimizes the placement and stabilization difficulties associated with the goniometer, there is more associated cost and the inclinometer is unable to measure hip range of motion in the horizontal plane [2,4]. Recent generations of smartphones have emerged as an alternative instrument for the measurement of joint ROM through the development of applications that utilize the smartphone's accelerometer and magnetometer. The potential clinical use of a smartphone has led to research investigating the validity and reliability of smartphone applications that measure joint ROM, focusing primarily on the spine, shoulder, knee and ankle [8–14].

Findings of this research indicate that smartphone applications for the measurement of joint range of motion have good intra-rater reliability, poor to good inter-rater reliability, and moderate to good inter-instrument validity with a goniometer or inclinometer. Research regarding the use of smartphone applications for the measurement of hip range of motion, however, is limited.

To date, only one study has investigated the reliability and validity of a smartphone application to measure hip ROM [15]. Charlton et al. compared the smartphone and a bubble inclinometer to a three dimensional motion analysis system (3DMA) in the measurement of hip flexion, abduction, adduction, internal rotation (IR) and external rotation (ER) of healthy adult males. The authors report good to excellent intra-rater reliability (ICC>0.75) for the measurement of hip flexion, supine IR, supine ER, and sitting IR and moderate reliability (ICC=0.63-0.68) for abduction, adduction, and sitting ER. In addition, the smartphone intra-rater reliability was comparable to that of the bubble inclinometer. Although inter-instrument validity of the smartphone against the 3DMA was supported in all hip motions except supine ER, there was fixed or proportional bias in three motions.

Research regarding the reliability and validity of clinical instruments used to measure hip ROM is limited. While acceptable intra-rater reliability of hip ROM measurements has been reported for both the goniometer and inclinometer, the inter-rater reliability of the devices are not as well supported and inter-instrument validity results do not support the interchangeable use of the two devices when measuring hip ROM [2,7,15,16]. Furthermore, the goniometer was found to overestimate hip PROM as compared to an electromagnetic tracking system [16]. The results reported by Charlton et al show promise in the use of a smartphone application to provide reliable and valid hip ROM measurements. Additional research is needed, however, to examine smartphone inter-rater reliability and to compare the Smartphone to the goniometer.

The purpose of this study was to establish the inter-rater and intra-rater reliability of smartphone measurements of hip flexion, extension, abduction, adduction, IR and ER. Additionally, the inter-instrument validity of the smartphone against the goniometer for each of the aforementioned measurements was assessed.

Methods

Thirty participants were recruited as a sample of convenience (11 male and 19 female; age range 20-40 years; mean age 25.6±3.5 years). A priori power analysis indicated that to achieve a power of .80 with P<0.05 and a minimally significant intraclass correlation coefficient (ICC) value of 0.75, a sample size of 26 participants would be required. Participants were excluded from this study if they had any present or past hip or spine pathology and had acute pain in the hips or low back. All participants were informed of the purpose of the study and signed an informed consent document prior to data collection. The study was approved by the Institutional Review Board at Western Kentucky University.

Three second year doctoral physical therapist (DPT) students collected data for this study. Two students served as raters and measured hip range of motion with the two devices while the third student served as an observer and recorded measurement data.

A 30.5 cm, 360°goniometer, marked in 1°increments (Sammons-Preston/Rolyan) was used to measure hip range of motion (Figure 1). A smartphone application (3D Protractor, V 2.1) downloaded for free from the App Store for iPhone was used to measure hip range of motion (Figure 2). The smartphone used in this study was an Apple iPhone®5. Raters measuring hip PROM were blinded to all measurements obtained. The face of the standard goniometer and the screen of the smartphone was covered by paper to prevent the raters performing hip ROM testing from viewing the measurement values. Only the top right corner of the smartphone screen was visible to the raters to allow access to the start/stop button to operate the application.

An apparatus was constructed out of PVC pipe material to serve as a target bar for the end point of each hip motion (Figure 3). For hip extension, a mobilization wedge was used to passively support the thigh in a position of hip extension. Standardized end positions of each hip motion were utilized to allow the identification of measurement error with each device that is independent of participant factors and rater skills in determining end ROM during

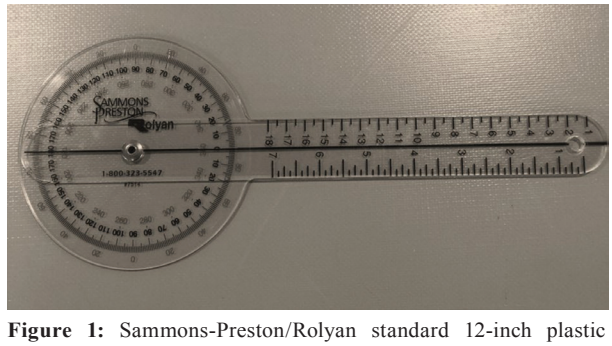

Figure 1: Sammons-Preston/Rolyan standard 12-inch plastic goniometer (model 2936770).

Figure 2: Screenshot of the 3D Protractor, V 2.1 on iPhone.

repeated movements of each hip motion examined. This allowed measurement reliability to be examined based on two factors, instrument related error and rater skill in use of each instrument. The use of the apparatus to standardize the end position of hip ROM also minimized the potential of a learning effect during repeated trails of a hip motion, which eliminated the need to randomize the order in which hip measurements were performed.

After a standardized warm-up, the participant's LE was measured and marked with a pen for Smartphone placement. One line was drawn 5 cm above the suprapatellar pole for positioning of the Smartphone on the anterior thigh during the measurements of hip flexion, abduction and adduction. A second line was drawn 5 cm above the popliteal crease for Smartphone positioning on the posterior thigh to measure hip extension. A third line was drawn 10 cm below the tibial tuberosity for Smartphone positioning on the anterior lower leg to measure hip IR and ER. All hip ROM measurements were performed on the right hip only. Six motions at the hip were measured in the following order: flexion, extension, internal rotation (IR), external rotation (ER), abduction, and adduction.

Standardized procedures and positions were followed to align the goniometer and to position the participant for each hip measurement[17]. The supine position was used to measure hip flexion, abduction and adduction. Hip extension was measured in prone with the knee extended. Hip IR and ER were measured in sitting with the knee in 90° flexion. For all motions, the hip was positioned in neutral specific to the plane of motion to be measured. During hip PROM, the rater stabilized the pelvis and adjacent joints to prevent movement into planes of motion not being measured.

General procedures for the measurement of hip ROM commenced placing the participant in supine, prone or sitting, positioning the apparatus to establish hip motion endpoint, and aligning the LE in neutral starting position. The smartphone was placed on the participant's LE and the "start" button on the smartphone application was activated by rater A. The observer recorded the initial hip angle from the smartphone. After removing the smartphone, rater A aligned the goniometer according to ROM procedures for each hip motion and measured

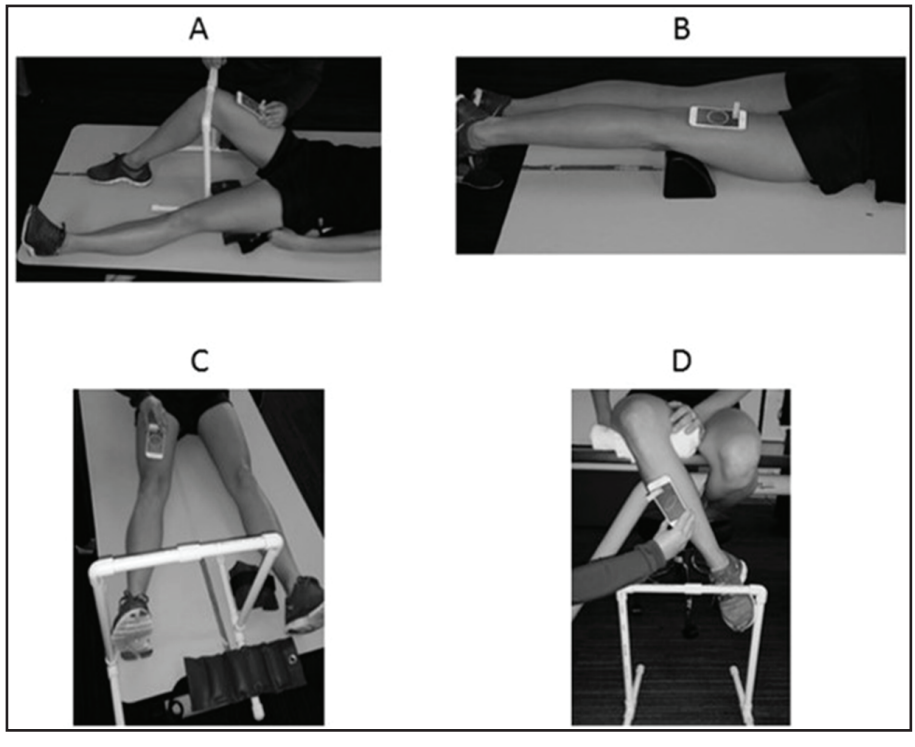

Figure 3: Testing position and apparatus used to designate end position during hip passive range of motion measurements: A=Hip flexion; B=Hip extension; C=Abduction and adduction; D=Internal and external rotation

the initial hip angle which was documented by the observer. Rater A then passively moved the participant's hip through the specified motion to the end point marked by the apparatus. While holding the LE at the end ROM, rater A placed the smartphone on the identified landmarks on the LE and activated the "stop" button on the smartphone to capture the final hip angle which was read and recorded by the observer. The smartphone was removed and the goniometer was aligned by Rater A to measure the endpoint ROM which was read and recorded by the observer. This procedure was repeated for all trials of a hip motion and for all hip motions. Rater A performed three trials of hip PROM measurements with the smartphone and the goniometer, followed by Rater B who repeated the measurement process with each device. The participant was allowed a rest period at any time during the measurement process upon request.

Data Analysis

Statistical analysis was carried out using SPSS version 21 (SPSS Inc., Chicago, IL, USA) for Windows. Descriptive statistics for measures of hip ROM taken by each investigator are reported using mean and standard deviation. Paired t-tests on the differences of measurements obtained between raters and between devices were used to ensure the absence of systematic bias. The level of significance was set at $p < .05$ for all statistical tests.

A two-way mixed model intraclass correlation coefficient ($ICC_{3,1}$) were used to analyze intra-rater reliability of 3 repeated measurements for each motion performed by each rater. Inter-rater reliability was examined using a two-factor random model 2 $ICC_{2,3}$, using the average of the three ROM measurements taken for each hip motion. A model 3 $ICC_{3,3}$ was used to analyze inter-instrument reliability (smartphone and goniometer), using the average of three ROM measurements taken for each hip motion. 95% confidence intervals (CI) were reported for all ICCs. Reference values used for the interpretation of ICC values for rater reliability and inter-instrument reliability were as follows: < 0.50, poor; 0.50-0.75, moderate; > 0.75 good [18].

Since ICCs may be influenced by high inter-subject variability and yield a large ICC despite poor trial to trial consistency of measures, absolute reliability was also assessed using standard error of measurement (SEM), and 95% limits of agreement (LOA). SEM was calculated using the formula: $SEM=SD\sqrt{(1-ICC)}$, where SD is the pooled standard deviations of all repeated measurements for a hip motion [18]. The 95% limits of agreement (LOA) between raters and devices were calculated using the formula 95% LOA=MD*2SDd where MD is the mean difference between raters or devices and SDd is the standard deviation of the difference scores between raters or devices for hip ROM measures [18]. SEM and LOA were rounded to the nearest degree to reflect the smallest until of measurement with the goniometer.

Results

Intra-Rater Reliability

Mean values for hip PROM for each rater are displayed in Table 1. Intra-rater reliability analysis including ICC with 95% CI and SEM are presented in Table 2. Intra-rater reliability was moderate to good for both raters using the Smartphone (ICC=0.47-0.99) and the goniometer (ICC=0.61-0.98). When comparing intra-rater reliability for each rater, Smartphone ICCs were good in 6 out of 6 motions for rater A and 4 out of 6 for rater B. Goniometric ICCs were lower, with 3 out of 6 motions having good reliability for each rater. For each instrument, intra-rater reliability was highest for the motions of flexion, IR and ER (ICC=0.95-0.99). SEM ranged from 1° to 3° for the smartphone and from 1° to 2° for the goniometer.

Inter-Rater Reliability

Inter-rater reliability analysis including mean difference between raters A and B, ICC with 95% CI, SEM and 95% LOA are presented in Table 3. Inter-rater reliability ICCs were moderate to good for measurements made with the Smartphone (ICC=0.70-0.99) and poor to good for measurements made with the goniometer (ICC=0.05-0.99). For both instruments, inter-rater reliability was good in 4 out of 6 motions, with the highest ICCs occurring for hip flexion, IR and ER (ICC=0.89-0.99). For all motions except hip abduction, there were no differences between raters in hip ROM when using the Smartphone ($p > 0.075$), indicating no systematic bias. All goniometric hip ROM values, however, were significantly lower for rater A, suggesting systematic bias between raters ($P<0.017$). SEM of smartphone inter-rater measurements ranged from 1° to 2°. SEM of goniometric hip measurements ranged from 1° to 3°. The 95% LOA suggest that hip ROM measurements of rater A may range from being 6° less to 5° greater than rater B when using the Smartphone and from -10° less to 8° higher than rater B when using the goniometer.

Table 1: Smartphone and goniometric hip range of motion measurements

	Rater A		Rater B	
	Smartphone Mean (SD)	**Goniometer Mean (SD)**	**Smartphone Mean (SD)**	**Goniometer Mean (SD)**
Flex	43.2 (7.6)	40.7 (5.5)	43.5 (6.8)	42.8 (5.7)
Ext	13.9 (2.3)	9.6 (1.9)	14.0 (2.1)	13.8 (2.1)
IR	27.0 (5.4)	26.0 (6.1)	27.0 (5.4)	27.9 (5.8)
ER	26.9 (5.1)	30.5 (5.9)	26.9 (5.2)	27.0 (6.1)
Abd	11.9 (2.2)	12.3 (2.6)	12.6 (2.0)	13.3 (2.2)
Add	19.5 (3.6)	11.2 (2.5)	20.1 (2.7)	12.2 (2.3)

SD: Standard deviation; Flex: Flexion; Ext: Extension; IR: Internal rotation; ER: External rotation; Abd: Abduction; Add: Adduction

Inter-Instrument Reliability

Inter-instrument reliability analysis including mean difference between instruments, ICCs with 95% CI, SEM, and 95% LOA are presented in Table 4. There were significant differences between devices for all measurements performed by rater A (P<0.01) except for

Table 2: Intra-rater reliability for hip passive range of motion

	Rater A Smartphone			Rater A Goniometer		
	ICC	**ICC 95% CI**	**SEM (°)**	**ICC**	**ICC 95% CI**	**SEM (°)**
Flex	0.99	0.98-0.99	1	0.96	0.93-0.98	1
Ext	0.75	0.53-0.87	1	0.7	0.45-0.85	1
IR	0.97	0.94-0.98	1	0.98	0.96-0.99	1
ER	0.96	0.93-0.98	1	0.98	0.97-0.99	1
Abd	0.68	0.41-0.83	2	0.69	0.43-0.84	2
Add	0.77	0.59-0.88	2	0.71	0.47-0.85	2
	Rater B Smartphone			**Rater B Goniometer**		
	ICC	**ICC 95% CI**	**SEM (°)**	**ICC**	**ICC 95% CI**	**SEM (°)**
Flex	0.99	0.98-0.99	1	0.95	0.91-0.97	1
Ext	0.82	0.67-0.91	1	0.7	0.45-0.85	1
IR	0.96	0.93-0.98	1	0.96	0.93-0.98	1
ER	0.97	0.94-0.98	1	0.98	0.97-0.99	1
Abd	0.59	0.26-0.79	2	0.64	0.34-0.82	2
Add	0.47	0.03-0.73	3	0.61	0.28-0.80	3

ICC: Intraclass correlation coefficient; 95% CI: 95% confidence intervals for ICC; SEM: Standard error of measurement; MDC_{95}: Minimal detectable change; Flex: Flexion; Ext: Extension; IR: Internal rotation; ER: External rotation; Abd: Abduction; Add: Adduction

Table 3: Inter-rater reliability for hip passive range of motion measurements

			Smartphone interrater reliability		
Variables	**Mean Diff (SD)**	**P value**	**ICC (95%CI)**	**SEM (°)**	**95% LOA**
Flex	-0.34 (1.3)	0.19	0.99 (.98-.99)	1	-3, 2
Ext	-0.13 (1.9)	0.72	0.77 (.51-.89)	1	-4, 4
IR	0.04 (2.3)	0.92	0.96 (.91-.98)	1	-5, 5
ER	0.03 (2.4)	0.96	0.95 (.89-.98)	1	-5, 5
Abd	-0.67 (1.9)	0.08	0.70 (.39-.86)	1	-5, 3
Add	-0.62 (2.9)	0.25	0.73 (.44-.87)	2	-6, 5
			Goniometric interrater reliability		
Variables	**Mean Diff (SD)**	**P value**	**ICC (95%CI)**	**SEM (°)**	**95% LOA**
Flex	-2.2 (2.3)	0.01	0.92 (.56-.97)	2	-7, 3
Ext	-4.2 (2.7)	0.01	0.05 (-.16-.31)	3	-10, 1
IR	-1.8 (2.4)	0.01	0.94 (.74-.98)	1	-7, 3
ER	3.5 (2.2)	0.01	0.89 (-.06-.97)	2	-1, 8
Abd	-1.0 (2.2)	0.02	0.72 (.34-.87)	1	-5, 3
Add	-0.93 (2.0)	0.02	0.76 (.47-.89)	1	-5, 3

Mean diff: Mean difference between raters (rater A – rater B; ICC: Intraclass correlation coefficient and 95% confidence intervals; SEM: Standard error of measurement; MDC_{95}: Minimal detectable change; 95% LOA: 95% limits of agreement

Table 4: Concurrent reliability analysis for hip passive range of motion measurements between instruments

			Rater A inter-instrument reliability		
Variables	**Mean Diff (SD)**	**Pvalue**	**ICC (95%CI)**	**SEM (°)**	**95% LOA (°)**
Flex	2.6 (3.0)	0.01	0.95 (.89-.97)	2	-4, 9
Ext	4.3 (2.7)	0.01	0.25 (-.57-.64)	3	-1, 10
IR	0.96 (2.0)	0.02	0.97 (.93-.99)	1	-3, 5
ER	-3.6 (2.6)	0.01	0.94 (.88-.97)	2	-9, 2
Abd	-0.41 (2.4)	0.36	0.67 (.31-.84)	1	-5, 4
Add	8.2 (2.7)	0.01	0.73 (.43-.87)	3	-3, 14
			Rater B inter-instrument reliability		
Variables	**Mean Diff (SD)**	**P value**	**ICC (95%CI)**	**SEM (°)**	**95% LOA (°)**
Flex	0.74 (3.3)	0.23	0.94 (.97-.97)	2	-6, 7
Ext	0.23 (2.2)	0.57	0.61 (.19-.82)	1	-4, 5
IR	-0.91 (2.9)	0.09	0.93 (.86-.97)	2	-7, 5
ER	-0.11 (3.3)	0.89	0.91 (.81-.96)	2	-7, 7
Abd	-0.76 (1.9)	0.04	0.74 (.44-.87)	2	-5, 3
Add	7.9 (2.4)	0.01	0.72 (.40-.87)	1	-3, 12

Mean diff: Mean difference between instruments (smartphone – goniometer); ICC: Intraclass correlation coefficient and 95% confidence intervals; SEM: Standard error of measurement; MDC_{95}: Minimal detectable change; 95% LOA: 95% limits of agreement

hip abduction. Specific patterns of differences between devices in ROM measurements conducted by rater A were not noted. For rater B, differences in ROM measurements between devices were non-significant except with hip adduction (P<0.04). Inter-instrument reliability ICCs were poor to good, ranging from 0.25- 0.97. For both raters, inter-instrument reliability was good for hip flexion, IR and ER (ICC=0.91-0.97) and moderate to poor for hip extension, abduction, adduction (ICC=0.25-0.74). The SEM for hip ROM measured by both devices were similar for both raters with SEM ranging from 1° to 3°. The 95% LOA suggest that the difference between hip ROM measurements performed with a Smartphone and goniometer varied between –9° to 10° for rater A and -7° to 12° for rater B.

Discussion

To our knowledge, this is the first study to evaluate intra-rater, inter-rater and inter-instrument reliability of a smartphone application for measuring hip PROM in all 6 motions. Although only one other study has examined reliability and validity of smartphone application to assess hip joint ROM, this study did not compare the smartphone to a goniometer, did not include assessment of inter-rater reliability or the motion of hip extension, and utilized only male participants [15]. Smartphone intra-rater, inter-rater and inter-instrument reliability was moderate to good (ICC>0.59) for all motions except hip adduction (intra-rater ICC=0.47) and hip extension (inter-instrument ICC=0.25). It is suggested that the ICC for reliability should exceed 0.90 for clinical measurements and that the clinical interpretation of reliability must involve the lower 95% CI [18]. Thus, a major finding in this study was that, when reviewing ICCs based on this suggestion and considering the lower bounds of 95% CI for all ICCs, smartphone intra-rater, inter-rater and inter-instrument reliability were >0.90 for the motions of hip flexion, IR and ER. A second major finding is that, for these motions, agreement of repeated measurements within raters, between raters and between instruments was also supported through low values in all measures of absolute measurement error.

Smartphone intra-rater reliability reported in the present study is higher than those reported by Charlton for motions performed in the same position (ICC=0.63-0.94) [15]. In addition, the present study found 3 out of 6 motions to exceed 0.90 while only one hip measurement, IR, resulted in ICC>0.90 in the aforementioned study. Charlton utilized end-range passive range of motion whereas the present study utilized a target to designate ending ROM [15]. As was intended in the design of the present study, the target may have provided greater consistency between repeated trails of a hip motion for each participant by minimizing participant related factors such as variations of movement performance that might occur as a learning effect or as mobility gains/losses from one trial to the next. This allowed rater reliability to be examined based on error inherent to the instrument and to the skill of the rater in use of the instrument. Although, both studies were similar in use of a smartphone to measure hip ROM, different applications were utilized. Charlton utilized a custom smartphone application, "Hip ROM Tester" which was designed by a co-author on the study, whereas the current study utilized 3D Protractor application [15]. Differences in functionality between the applications may have also contributed to the differences in intra-rater reliability between the two studies.

In the current study, the lowest intra-rater and inter-rater ICC values attained with both devices and by both raters were for the motions of hip extension, abduction and adduction. The lower ICCs in these motions as compared to hip flexion, IR and ER may relate to differences in soft tissue mass in areas where the instruments were placed and to the methods of performing the hip motions. During the measurement of hip extension, abduction and adduction, both the smartphone and the moving arm of the goniometer were aligned on the thigh as compared to the lower leg when measuring hip IR and ER. The greater muscle mass in the hip and thigh areas as compared to the lower leg may have contributed to greater difficulty when palpating landmarks to align the goniometer and instability in the orientation of the smartphone, thus compromising measurement reproducibility.

Both the present study and that of Charlton found low smartphone intra-rater reliability for hip abduction and adduction (ICC=0.68) [15]. Charlton attributes the low abduction and adduction intra-rater ICCs to the use of a side-lying position which may have reduced the stability of the participant. In the present study, smartphone measures of hip abduction and adduction were performed in supine according to standard goniometric procedures yet yielded moderate to poor intra-rater and inter-rater reliability. In addition to previously mentioned soft tissue related factors potentially lending to lower ICC for these motions, inadequate stabilization of the pelvis may also be a contributing factor. While the 3-D Protractor application is one of the few applications that are able to measure in the horizontal plane, it is sensitive to orientation of the smartphone. Any rotation of the hip during horizontal plane abduction and adduction would cause the smartphone to tilt, affecting the measurement value the measurement value recorded by the application and potentially contributing to measurement error which would be reflected as lower intra and inter-rater ICC values.

Although inter-rater relative reliability of smartphone measurements was greater than what is considered clinically meaningful (ICC >0.90) in only 3 out of 6 motions, absolute agreement between raters for all smartphone hip measurements was supported through low SEM (<2°), and 95% LOA ranging from -6° to 5°. The low SEM suggests that measurement error between raters would be less than 2°. While there are no previous studies to compare these results, these findings are similar to the recommendation of Boone et al. [20] who suggested a change of at least 6° is needed to document real change in lower extremity ROM when measured by more than one clinician.

The inter-rater reliability of goniometric hip measurements was found to be lower as compared to inter-rater reliability

of smartphone measurements in all motions except hip abduction and adduction (goniometric ICCs=0.05-0.94; smartphone ICCs=0.70-0.99). It is suggested that ICCs should also be interpreted with consideration of the lower 95% CI. [18] Thus, further examination of the lower 95% CI for ICC values show poor to moderate (<0.75) inter-rater reliability in all goniometric hip ROM measures. While the findings of poor to moderate inter-rater reliability of goniometric hip ROM are in agreement with other studies, the ICC values reported in the present study, with the exception of hip extension, are higher than values reported in these studies which ranged from 0.22-0.86 for healthy adults and from 0.53-0.73 in adults with hip osteoarthritis [2,6]. The higher ICC in the present study is most likely due to the purposeful use of a target end point for hip PROM to establish reliability that is independent of patient related factors as compared to use of full PROM in the prior studies.

Both raters attained inter-instrument ICC values that could be deemed clinically acceptable (≥0.90) for the motions of flexion, IR and ER, however, there were significant differences between the instruments for these motions assessed by rater A. When considering all 6 hip motions, there was inter-instrument bias (significant differences) in 5 out of 6 hip PROM measurements performed by rater A as compared to 1 out of 6 measurement performed by rater B (Table 4). Similar findings were reported in two prior studies that examined inter-instrument reliability between clinical instruments used to measure hip PROM [7,15]. Charlton found fixed or proportional bias between the Smartphone and 3DMA in 4 of 7 hip ROM measurements despite good ICC values[15]. In a study examining the inter-instrument validity of a digital inclinometer and universal goniometer, Roach et al. found significant differences between the goniometer and inclinometer when measuring PROM of hip extension, IR and ER, with mean differences ranging between 3-5°[7].

For the motions of hip flexion, IR and ER, low values of measurement error, 1°-2°, were obtained, providing evidence of agreement between the instruments. In addition, for these motions, the 95% LOA suggests that Smartphone measurements may range from being 7° less to 9° greater than goniometric measures. Since no prior studies have examined inter-instrument validity of Smartphone and goniometric hip PROM measurements, a comparison between this study and previous research cannot be made.

Reliability determination is relevant only within the context in which measurements are performed. Thus, the findings regarding reliability of Smartphone hip ROM measurements should only be generalized to young, healthy adults and procedures in which ROM is performed to a designated end position. Future research should examine the reliability of Smartphone hip ROM measurements using a symptomatic population. The use of full available hip ROM performed without a target end-point would allow expansion of the current research to include examination of participant related factors that may influence reliability. The design of the present study involved conducting all hip PROM measurements during one session within a time span of approximately 1 hour. While this might simulate the length of a treatment session where hip PROM may change as a result of interventions provided, caution should be used when applying these findings to between day comparisons of hip ROM. Raters selected for this study were inexperienced in use of both instruments to measure ROM beyond classroom instruction in goniometric procedures and training in the procedures outlined in the current study, thus the comparison of novice and experienced users of both instruments would expand the findings of the current study to simulate the varied experience level that is commonly found in a clinical setting. Lastly, inter-rater and inter-instrument reliability was assessed using the average of three repeated measurements. Clinical ROM measurements, however, are most often documented with a single trial. Future research should examine inter-rater and inter-instrument reliability of Smartphone hip PROM assessed with a single measurement.

Conclusion

These findings of this study support intra-rater reliability of both instruments when measuring hip PROM. Inter-rater reliability, however, was supported only for the smartphone due to the lack of absolute agreement between raters for all goniometric hip measurements. While measurements performed by rater B resulted moderate to good inter-instrument reliability and non-significant differences between Smartphone and goniometric hip PROM, there were significant differences between devices in measurements performed by rater A. Thus, caution should be used if the instruments are to be used interchangeably in order to quantify within session hip PROM.

Conflict of Interest: None declared.

References

1. Beneciuk JM, Bishop MD, George SZ. Clinical Prediction Rules for Physical Therapy Interventions: A Systematic Review. Phys Ther. 2009 Feb 1;89(2):114–24.
2. Prather H, Harris-Hayes M, Hunt D, Steger-May K, Mathew V, Clohisy J. Hip range of motion and provocative physical examination tests reliability and agreement in asymptomatic volunteers. PM R. 2010 Oct;2(10):888–95.
3. Reinhold Ganz TS. The lesser trochanter as a cause of hip impingement: Pathophysiology and treatment options. Hip Int J Clin Exp Res Hip Pathol Ther. 2013;
4. Bierma-Zeinstra SM, Bohnen AM, Ramlal R, Ridderikhoff J, Verhaar JA, Prins A. Comparison between two devices for measuring hip joint motions. Clin Rehabil. 1998 Dec;12(6):497–505.
5. Herrero P, Carrera P, García E, Gómez-Trullén EM, Oliván-Blázquez B. Reliability of goniometric measurements in children with cerebral palsy: A comparative analysis of universal goniometer and electronic inclinometer. A pilot study. BMC Musculoskelet Disord. 2011 Jul 10;12:155.
6. Poulsen E, Christensen HW, Penny JØ, Overgaard S, Vach W, Hartvigsen J. Reproducibility of range of motion and muscle strength measurements in patients with hip osteoarthritis – an inter-rater study. BMC Musculoskelet Disord. 2012 Dec 6;13:242.
7. Roach S, San Juan JG, Suprak DN, Lyda M. Concurrent validity of digital inclinometer and universal goniometer in assessing passive hip mobility in healthy subjects. Int J Sports Phys Ther. 2013 Oct;8(5):680–8.

8. Anderson D Stephen, English E Alton, Varee K Bai, Diaz D, Lowe D. Reliability and validity of an iPhone® inclinometer as compared to a universal goniometer as a tool for measuring joint motion of the shoulder in apparently healthy subjects. Orthop Phys Ther Pract. 2013 Mar;25(1):34–8.
9. Kolber MJ, Pizzini M, Robinson A, Yanez D, Hanney WJ. The reliability and concurrent validity of measurements used to quantify lumbar spine mobility: an analysis of an iphone® application and gravity based inclinometry. Int J Sports Phys Ther. 2013 Apr;8(2):129–37.
10. Tousignant-Laflamme Y, Boutin N, Dion A, Vallée C-A. Reliability and criterion validity of two applications of the iPhone™ to measure cervical range of motion in healthy participants. J Neuroeng Rehabil. 2013;10(1):69.
11. Jenny J-Y. Measurement of the knee flexion angle with a Smartphone-application is precise and accurate. J Arthroplasty. 2013 May;28(5):784–7.
12. Johnson LB, Sumner S, Duong T, Yan P, Bajcsy R, Abresch RT, de Bie E, Han JJ. Validity and reliability of smartphone magnetometer-based goniometer evaluation of shoulder abduction - A pilot study. Man Ther. 2015 Mar 17;
13. Vohralik SL, Bowen AR, Burns J, Hiller CE, Nightingale EJ. Reliability and validity of a smartphone app to measure joint range. Am J Phys Med Rehabil Assoc Acad Physiatr. 2015 Apr;94(4):325–30.
14. Shin S Han, Ro D Hyun, Lee O-S, Oh J Han, Kim S Hoon. Within-day reliability of shoulder range of motion measurement with a smartphone. Man Ther. 2012 Aug;17(4):298–304.
15. Charlton PC, Mentiplay BF, Pua Y-H, Clark RA. Reliability and concurrent validity of a Smartphone, bubble inclinometer and motion analysis system for measurement of hip joint range of motion. J Sci Med Sport. 2015 May 1;18(3):262–7.
16. Nussbaumer S, Leunig M, Glatthorn JF, Stauffacher S, Gerber H, Maffiuletti NA. Validity and test-retest reliability of manual goniometers for measuring passive hip range of motion in femoroacetabular impingement patients. BMC Musculoskelet Disord. 2010;11:194.
17. Bandy PhD PT SCS ATC WD, Reese PhD PT NB. Joint Range of Motion and Muscle Length Testing, 2e [Internet]. [cited 2015 May 19]. Available from: http://www.biblio.com/book/joint-range-motion-muscle-length-testing/d/817100305
18. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Upper Saddle River, NJ: Pearson Prentice Hall; 2009.
19. Bland JM, Altman DG. Measurement error. BMJ. 1996 Sep 21;313(7059):744.
20. Boone DC, Azen SP, Lin CM, Spence C, Baron C, Lee L. Reliability of goniometric measurements. Phys Ther. 1978 Nov;58(11):1355–60.