Concurrent validity and reliability of the Simple Goniometer iPhone® app compared with the Universal Goniometer

Authors

Dr Anne Jones, PhD. Senior Lecturer, Physiotherapy, James Cook University, Townsville Qld 4811, Australia. anne.jones@jcu.edu.au

Dr Rebecca Sealey, PhD. Senior Lecturer, Institute of Sport and Exercise Science, James Cook University, Townsville Qld 4811, Australia. rebecca.sealey@jcu.edu.au

Dr Michael Crowe, PhD. Research Officer, School of Public Health, Tropical Medicine and Rehabilitation Sciences, James Cook University, Townsville Qld 4811, Australia. michael.crowe@my.jcu.edu.au

Dr Susan Gordon, PhD. Associate Professor, Physiotherapy, James Cook University, Townsville Qld 4811, Australia. susan.gordon2@jcu.edu.au

ABSTRACT

Objectives: The aim of this study was to assess the concurrent validity and reliability of the Simple Goniometer iPhone® app compared to the Universal Goniometer.

Design: Within subject comparison design comparing the Universal Goniometer with the Simple Goniometer app.

Setting: James Cook University, Townsville, Queensland, Australia.

Participants: Thirty-six volunteer participants, with a mean age of 60.6 years (SD 6.2).

Interventions: Not applicable.

Main Outcome measure: Thirty-six participants performed three standing lunges during which the knee joint angle was measured with the Simple Goniometer app and the Universal Goniometer.

Results: There were no significant differences in the measures of individual knee joint angles between the Universal Goniometer and the Simple Goniometer app. Pearson correlations of 0.96–0.98 and intraclass correlation coefficients of 0.97–0.99 (95% CI 0.95–1.00) were recorded for all measures. Using the Bland-Altman method, the standard error of the mean of the differences and the standard deviation of the mean of the differences were low.

Conclusion: The measurements from the Simple Goniometer iPhone® app were reliable and possessed concurrent validity for this sample and protocol when compared to the Universal Goniometer.

Key words: goniometry, joint range, measurement, knee

List of abbreviations: Universal Goniometer (UG), Micro-electromechanical system (MEMS), Simple Goniometer (SG), Intraclass correlation coefficients (ICC)

INTRODUCTION

Healthcare professionals use goniometry to objectively measure joint range of movement. For over 50 years the full circle, manual Universal Goniometer (UG) (Chattanooga, CA, USA) has been the most widely used tool for measuring joint range of motion (Gajdosik & Bohannon, 1987). To ensure reliable measurement, standardised, specific positions and landmarks are used to measure each joint movement (Clarkson, 2000; Norkin & White, 2003). With the introduction of smartphones, applications for measuring joint range of motion have been developed. This method of measuring joint range of motion is new, readily available and inexpensive (excluding the cost of the smartphone), however the validity and reliability of many of these tools have not been established.

Previous research has been undertaken to determine the reliability and validity of the UG. There is high intra-rater reliability for UG measurement of knee extension and flexion (Pearson's $r \geq 0.91$) (Gogia, Braatz, Rose, & Norton, 1987; Rothstein, Miller, & Roettger, 1983) and high inter-rater reliability for knee flexion (Pearson's $r \geq 0.82$) (Rothstein, et al., 1983). Goniometry is also a sufficiently valid measurement tool when compared to measuring knee joint angle via radiography (Pearson's $r \geq 0.97$, ICC 0.98–0.99) (Gogia, et al., 1987). However, in relation to the first 15 degrees of knee flexion there is a significant difference ($p < 0.01$) between goniometry and radiological measurement such that goniometry may be inaccurate (Enwemeka, 1986).

Knee angles can also be measured with high reliability using digital inclinometers (inter-rater, Pearson's r 0.83–0.94) (Cleffken, van Breukelen, Brink, van Mameren, & Damink, 2007). dos Santos, et al. (2012) found that for knee extension the digital inclinometer had greater intra- and inter- rater reliability (ICC 0.69–0.89) than the UG (ICC 0.46–0.55) but the digital inclinometer and UG had similar reliability for knee extension (ICC 0.91–0.98).. Even though digital inclinometers are more accurate than the UG, they are more costly and, therefore, may not be available to all practitioners (Shin, Ro, Lee, Oh, & Kim, 2012).

The inclusion of micro-electromechanical system (MEMS) accelerometers in smartphones means that applications are available which may be able to measure joint angles comparable with a UG (Shaw, Adames, Izzatt, Licina, & Askin, 2011). A strong correlation has been reported for scores obtained by a smartphone accelerometer and a UG (Pearson's r 0.79–0.97) for shoulder range of movement intra-rater reliability (smartphone ICC 0.90–0.99, UG ICC 0.80–0.99) and inter-rater reliability (smartphone ICC 0.63–0.90; UG ICC 0.64–0.91) (Shin, et al., 2012). The iGoniometer, a smartphone application software (app) for knee flexion, was used in a study to measure maximal active knee flexion angle with the result of a statistically significant smaller knee angle measurement of 1.4 degrees compared to a UG, despite a Pearson's correlation of $r = 0.93$ (Hambly, Sibley, & Ockendon, 2012). Although the difference was statistically significant, it is not likely to be clinically relevant given the smallest detectable differences reported by Cleffken, et al. (2007) was 7.0–9.4 degrees for active maximum knee flexion and 7.5–9.5 degrees for active range of movement.

The Simple Goniometer app (SG) for Apple iOS (http://www.ockendon.net/) measures joint angles by comparing the inclination of the segment above and below the joint. It is currently a paid app (US$0.99) compatible with the Apple iPhone®, iPod touch and iPad (Cupertino, CA, USA), and requires a minimum of iOS 4.0. Unlike the iGoniometer smartphone application, the SG has not yet been tested for reliability or validity.

The aim of this study was to investigate the concurrent validity and reliability of the SG compared with a UG for assessing knee joint angle in the lunge position. This is the first known study to consider the SG's use in a functional, weight bearing position.

METHODS

This study was undertaken during the pre-intervention phase of a whole body vibration study (Jones, Gordon, & Sealey, 2013). Consequently, the participants should be considered a convenience sample. Although participants were purposively recruited for the main study, they were included in this study because they were already present. The only criterion for this study was that the participant could maintain the lunge position long enough for both measurements to be taken. The age, inclusion and exclusion criteria, and other conditions, although reported here for completeness do not directly bear on this study's aim of testing SG score concurrent validity and reliability. Intra-rater reliability could not be tested because predetermined knee joint positions would be required and this would be in conflict with the main study. Furthermore, there was only one rater and inter-rater reliability could not be tested. Mentions of validity and reliability are in reference to the SG and UG only rather than the person using the instruments.

Ethical permission for this study was granted by the James Cook University Human Research Ethics Committee (H3976).

Recruitment

For the main study, participants' were required to be over the age of 50 and were recruited via an email to the staff of James Cook University who were encouraged to distribute the information to the broader community of Townsville, Australia. The inclusion and exclusion criteria were based on broad medical issues that may be aggravated by participation in the main study's vibration activities and, therefore, were a requirement, although not necessary, for the concurrent research project (Jones, et al., 2013). Inclusion criteria were that participants spoke English, were of good health, had no physical injury or limitation which would prevent them from participating in 10 minutes of vibration and were able to comfortably maintain a gentle lunge position. Exclusion criteria included severe cardiovascular disease, acute thrombosis, metal implants (including joint replacements), a pacemaker, severe diabetes, infection in the previous four weeks, severe migraines, malignant tumours, stents implanted in the previous four weeks, epilepsy, gallbladder or kidney stones, recent wounds in the previous four weeks, acute inflammation or pain, or cognitive deficits. All participation criteria were assessed through a pre-enrolment questionnaire.

Eligible participants were provided with an appointment during November and December 2011. All assessment sessions were undertaken at James Cook University in Townsville Australia. On arrival each participant was given an information sheet and informed consent form, which was completed and signed by participants prior to participation.

Research assistants

Two assistants collected the data. Assistant one was a registered physiotherapist experienced in using the UG and performed all knee measurements using the SG and UG. Assistant two was an exercise physiologist and was trained in the protocol for the study. They were responsible for recruitment, obtaining informed consent, participant safety, and recording measurements. Using two different assistants reduced the possibility of bias because assistant one was blinded from the scores obtained while using the SG and UG.

Instrumentation

A single UG which had a plastic 360 degree goniometer face and 10 inch movable arms and a single iPhone® 3GS (iOS 4.3.5) with the SG app were used for all measurements.

Procedure

Participants were asked to remove their shoes and the landmarks used for the placement of the iPhone® and UG were marked on the participants' skin by assistant one. Landmarks for the iPhone® were 10 cm above the superior border of the patella and 10 cm below the inferior pole of the patella (Figure 1). Landmarks for the UG were: fixed arm pointing to the greater trochanter of the femur; axis at the lateral epicondyle of the femur; and moving arm pointing towards the lateral malleolus of the fibula (Clarkson, 2000; Norkin & White, 2003).

Insert here **Figure 1 SG procedure**

One surface of the UG face was covered so that the numbers and marks could not been seen by assistant one. The UG was placed on the landmarks and when assistant one was satisfied with the angle of measure, they passed the goniometer to assistant two who read and recorded the joint angle.

For the SG, the onscreen goniometer and the angle were covered so that assistant one could not read the measured angle. Assistant one placed the edge of the iPhone® (that is, the opposite edge from the mute and volume buttons) against the participants' skin with the bottom end of the iPhone® 10 cm above the superior border of the patella (Figure 1, Step 1). The set button on the screen was pressed to record that position (Figure 1, Step1A screenshot). At this point the set button became the hold button and the angle was set to 0 degrees automatically by the SG app (Figure 1, Step 1B screenshot). The edge of the iPhone® was then placed against the participants' skin with the top end of the iPhone® 10 cm below the inferior pole of the patella and the hold button on the screen was pressed (Figure 1, Step 2) at which point the SG app calculated the angle automatically. The iPhone® was passed immediately to assistant two who logged the measurement as calculated by the

SG app. Assistant two pressed the reset button on the SG app (Figure 1, Step2 screenshot) clearing any information from the screen and returning the screen to that shown in Figure 1, Step1A. No other procedure or calibration was required for the SG app.

A participant's dominant leg was identified by asking, "Which leg do you use to kick a ball?" Participants were then positioned with their dominant foot a standardised 5 cm from a wall. They were then asked to gently lunge forward with their dominant leg so that their knee was as close as possible to the wall. The angle of the knee was measured using the UG and the SG as per the procedure outlined above. Participants were asked to remember this target angle.

Participants were moved away from the wall position to remove visual cues and were asked to gently lunge forward to reproduce the target angle. The knee joint angle was measured with the SG and UG, and results were recorded as measurement 1. The participant then stood upright for five seconds before repeating the gentle lunge to reproduce the target angle two more times, creating measurement 2 and measurement 3. Taking three measurements, like this, is a common procedure when ascertaining joint angle. The order of UG or SG to measure the knee joint angle was randomly selected by picking a card out of a box for each of the three measurements.

Data analysis

Sample size was calculated using the formulae for reliability studies based on confidence intervals described by Streiner and Norman (2008, pp. 198-202). With the number of instruments (k) equal to 2, the confidence interval around r (the reliability coefficient) of 0.05, and an estimated r of 0.95, the sample size (n) was calculated to be 25 participants. Pearson's correlation is a good measure for testing concurrent validity and was calculated for comparison with earlier goniometry studies. Intraclass correlation coefficients (ICC) for consistency (also known as relative agreement), is a good test for validity and a fair test for reliability, while ICC for absolute agreement is a good test for reliability and validity. Therefore, only ICCs for absolute agreement are reported. ICCs were calculated using the mixed effects model (also known as model 3, with the goniometers fixed effects and participants random effects). These tests were calculated using SPSS version 18.0.3 (Chicago, IL, USA). In addition, Bland-Altman plots, which represent the differences between the SG and UG in graph form, were produced using MedCalc version 12.3 (Mariakerke, Belgium). Scores for each individual measurement (1–3) are reported rather than mean measures because a mean may hide inconsistent results.

RESULTS

The study recruited 36 participants with a mean age of 60.6 years (SD 6.2). Ten (28%) participants were between 50 and 59 years of age and 26 (72%) participants were over 60 years of age. There were eight (22%) male and 28 (78%) female participants.

Measurements 1–3 for the UG and SG were found to be normally distributed. Overall, the UG measured a minimum knee joint angle of 35 degrees and a maximum of 85 degrees (range 50 degrees), and the SG measured a minimum knee joint angle of 36 degrees and a maximum of 89 degrees (range 53 degrees). Pearson's *r* was 0.96–0.98 for measurements 1–3 ($p < 0.001$, 95% confidence interval 0.92–0.99).

Using the Bland-Altman method, the SG measured +0.5 degrees more than the UG on average (Table 1). The standard error of the mean of the differences and the standard deviation of the mean of the differences were low (SD_{ME}, was 2.6 for measurement 1, 3.3 for measurement 2, and 2.3 for measurement 3; SE_{MD}, was 0.4 for measurement 1 and 3, and 0.6 for measurement 2). Figure 2 (A–C) illustrates the Bland-Altman plots for each measurement.

There was no indication of bias at the 95% confidence interval (CI) of limits of agreement. However, Participant 20 was a consistent outliner in measurements 1–3 where the mean difference between the UG and SG was –10 degrees. Participant 2 was an outlier of +6 degrees for measurement 1 (Figure 2A) and +12 degrees for measurement 2, but not for measurement 3.

The single measures ICC is the most appropriate coefficient because in practice only one goniometer would be used. However, the average measures are also shown in Table 2 because these ICCs are commonly reported in other research. The ICC(3,1) model for absolute agreement with n = 36 were: 0.97 (95% CI 0.95–0.99) for measurement 1; 0.96 (95% CI 0.92–0.98) for measurement 2; and 0.98 (95% CI 0.96–0.99) for measurement 3.

To determine whether the larger sample size of n = 36 had an effect on the ICCs, five random samples of 25 participants, the original sample size, were taken using the RANDOM command in SPSS. In each case, Participant 2 and 20 were included in the random samples. The most conservative of these five samples are in Table 3, which shows that for n = 25 the ICC(3,1) for absolute agreement were: 0.95 (95% CI 0.90–0.98) for measurement 1; 0.93 (95% CI 0.84–0.97) for measurement 2; and 0.97 (95% CI 0.92–0.98) for measurement 3.

Insert here **Table 1.** Bland-Altman method (measurements in degrees)

Insert here **Figure 2.** Bland-Altman plot, measurement 1, 2, and 3

Insert here **Table 2.** Intraclass correlation coefficients

Insert here **Table 3.** Conservative ICCs

DISCUSSION

There was a positive correlation ($r \geq 0.96$) for knee joint angle measurement using the UG and SG, thereby showing that scores obtained from the SG are concurrently valid when compared to the UG in this study. Even when using the most conservative calculations, ICCs were ≥ 0.93 between the instruments and the results were similar to the study by Gogia et al (Gogia, et al., 1987). The Bland-Altman method showed no bias between the instruments and the standard error and standard deviation of the mean differences between the instruments was low. The consistent +0.5 degree difference between the UG and SG could be due to measurer or instrument error. Overall, however, the scores obtained from the SG were reliable in this study.

In contrast to previous studies a functional, active, weight-bearing knee joint position was measured in this study. Active, non-weight bearing range of movement measurements have less variability and more reliability than passive range of movement measurements (Gajdosik & Bohannon, 1987). Thus the use of active movement in this study is likely to have negatively affected the reliability scores obtained using the UG and SG. Furthermore, Participant 2 and 20 were outliers and further analysis showed that both participants were measured with the UG first and then the SG. The variability in the measurements could be due to the participant fatigue between the UG and the SG measurements potentially reducing reliability.

The difficulty in locating landmarks when assessing the lower limb has been cited as a reason for lower reliability scores for the UG (Gajdosik & Bohannon, 1987). In functional positions, therefore, the SG may be easier to use for lower limb range of movement measurement because landmark identification is not required and less experienced clinicians may be inclined to use the SG for this reason. Conversely there is an increased cost associated with purchasing a smartphone and the necessary app compared with the UG.

This study investigated measurement of knee joint flexion between 35 and 89 degrees and with the iPhone® placed 10 cm above and below the knee joint. It is recommended that further studies are undertaken to confirm the findings of this study in knee joint angles that are less than 35 degrees and over 90 degrees and that assess different iPhone® placement positions to determine whether reliability is affected by joint angle or iPhone® placement.

Limitations

Participants of this study were over 50 years old with no joint replacements, acute pain or infection. Further studies need to be undertaken to determine if these results are applicable to people under the age of 50 or those with joint replacements or pain.

Intra-rater and inter-rater reliability testing were not undertaken in this study. Therefore, it is not possible to comment on whether the measurements from the SG have intra-rater or inter-rater reliability. Clinicians should be wary in using the SG or other similar apps without these reliability scores being available.

CONCLUSIONS

The scores obtained from the SG app for iPhone® showed that there was concurrent validity and reliability for knee joint angle as compared with the UG, as undertaken in this study. Further studies are required to investigate SG in the standard goniometry measuring positions, intra-rater reliability, inter-rater reliability, and a wider range of knee joint angles.

Acknowledgement

There are no funding sources to declare for this project. We would like to acknowledge Ms Fiona Pringle and Ms Lauren Clark for their assistance in recruitment and data collection for this study.

Declaration of interest

There are no declarations of interest for this study.

REFERENCES

Clarkson, H. (2000). *Musculoskeletal assessment. Joint range of motion and manual muscle strength* (2nd ed.). Philadelphia: Lippincott Williams and Wilkins.

Cleffken, B., van Breukelen, G., Brink, P., van Mameren, H., & Damink, S. O. (2007). Digitial goniometric measurement of knee joint motion. Evaluation of usefulness for research settings and clinical practice. *The Knee, 14*(5), 385-389. doi:10.1016/j.knee.2007.07.004

dos Santos, C. M., Ferreira, G., Malacco, P. L., Sabino, G. S., de Souza Moraes, G. F., & Felicio, D. C. (2012). Intra and inter examiner reliability and measurement error of goniometer and digital inclinometer use. *Revista Brasileira de Medicina do Esporte, 18*(1), 38-41. doi:10.1590/S1517-86922012000100008

Enwemeka, C. (1986). Radiographic verification of knee goniometry. *Scandinavian Journal of Rehabilitation Medicine, 18*(2), 47-49.

Gajdosik, R., & Bohannon, R. (1987). Clinical measurement of range of motion: review of goniomtery emphasizing reliabilty and validity. *Physical Therapy, 67*(12), 1867-1872.

Gogia, P., Braatz, J., Rose, S., & Norton, B. (1987). Reliability and validity of goniometric measurements at the knee. *Physical Therapy, 67*(2), 192-195.

Hambly, K., Sibley, R., & Ockendon, M. L. (2012). Level of agreement between a novel smartphone application and a long arm goniometer for the assessment of maximum active knee flexion by an inexperienced tester. *International Journal of Physiotherapy and Rehabilitation, 2*. Retrieved from http://ijptr.com

Jones, A., Gordon, S., & Sealey, R. (2013). The effect of a single session of Whole Body Vibration on mobility, lower limb joint position sense and balance of community dwelling older adults. *Internet Journal of Allied Health Science and Practice, 11*(3). Retrieved from http://ijahsp.nova.edu/

Norkin, C., & White, D. (2003). *Measurement of joint motion. A guide to goniometry* (3rd ed.). Philadelphia: FA Davis Company.

Rothstein, J., Miller, P., & Roettger, R. (1983). Goniometric reliabilty in a clinical setting. *Physical Therapy, 63*(10), 1611-1615.

Shaw, M., Adames, C., Izzatt, M., Licina, P., & Askin, G. (2011). Use of the iPhones for Cobb angle measurement in scoliosis. *European Spine Journal, Online First*. doi:10.1007/s00586-011-2059-0

Shin, S. H., Ro, D. H., Lee, O., Oh, J. H., & Kim, S. H. (2012). Within-day reliability of shoulder range of motion measurement with a smartphone. *Manual Therapy, 17*(4), 298-304. doi:10.1016/j.math.2012.02.010

Streiner, D., & Norman, G. (2008). *Health measurement scales: a practical guide to their development and their use* (4th ed.). Oxford: Oxford University Press.

Table 1 Bland-Altman method (in degrees)

n = 36	**MD**	**SE_{MD}**	**CI MD**	**SD_{MD}**	**CI LA**
Measurement 1	0.5	0.4	-0.4 – 1.4	2.6	-4.6 – 5.6
Measurement 2	0.5	0.6	-0.6 – 1.7	3.3	-6.0 – 7.1
Measurement 3	0.5	0.4	-0.3 – 1.3	2.3	-4.0 – 5.1

MD, mean of differences; SD_{MD}, Standard deviation of MD;
SE_{MD}, Standard error of MD; CI MD, 95% confidence interval of MD;
CI LA, 95% confidence interval of limits of agreement

Table 2 Intraclass correlation coefficients [ICC(3,1)] for absolute agreement (n = 36)

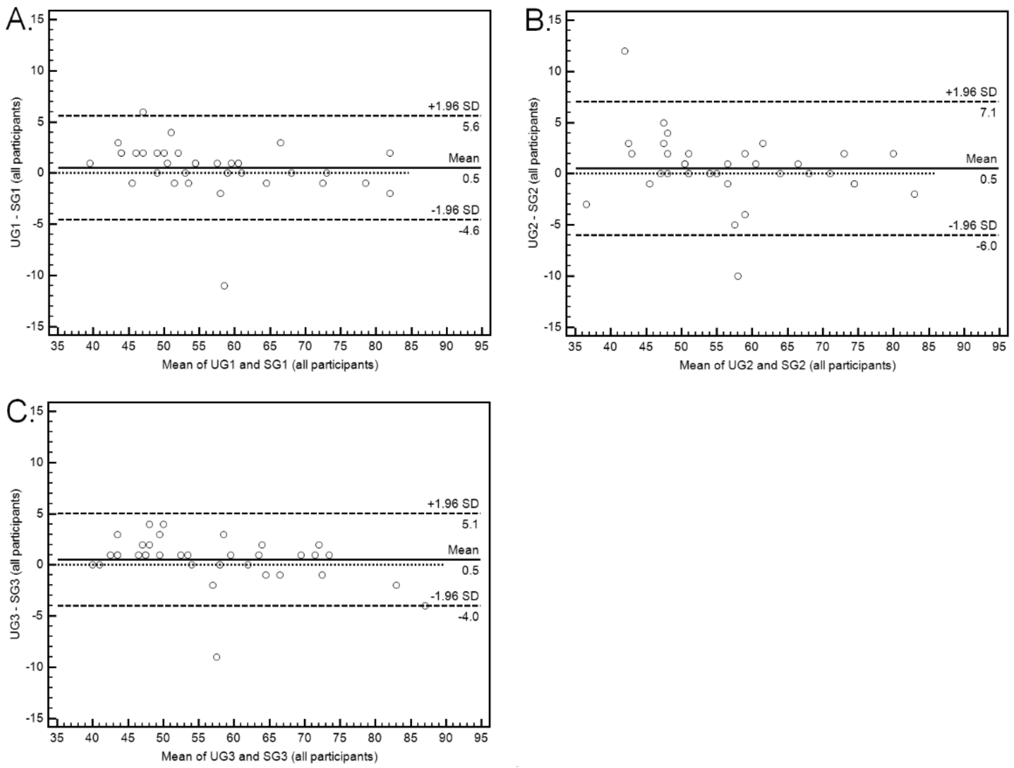
	Single measures		**Average measures**	
	A	(95% CI)	A	(95% CI)
Measurement 1	0.97	(0.95–0.99)	0.99	(0.97–0.99)
Measurement 2	0.96	(0.92–0.98)	0.98	(0.96–0.99)
Measurement 3	0.98	(0.96–0.99)	0.99	(0.98–1.00)

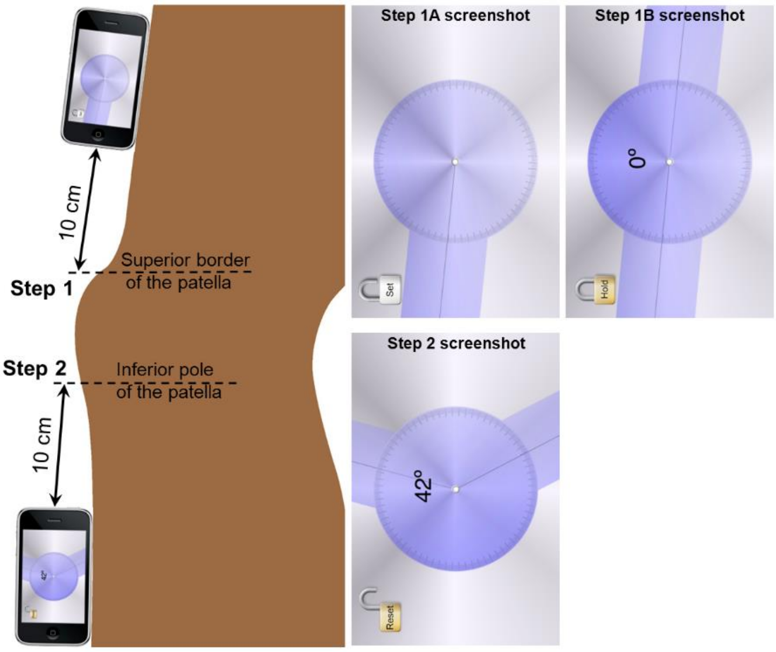
A, Absolute agreement; CI, confidence interval

Table 3 Conservative ICC(3,1) for absolute agreement (random selection of 25 participants, including Participant 2 and 20)

	Single measures		**Average measures**	
	A	(95% CI)	A	(95% CI)
Measurement 1	0.95	(0.90–0.98)	0.98	(0.95–0.99)
Measurement 2	0.93	(0.85–0.97)	0.96	(0.92–0.98)
Measurement 3	0.97	(0.92–0.98)	0.98	(0.96–0.99)

A, Absolute agreement; CI, confidence interval





Step 1A screenshot
Step 1B screenshot
Step 2 screenshot
10 cm
Superior border
of the patella
Step 1
Step 2
Inferior pole
of the patella
10 cm
0°
42°
Set
Hold
Reset