Validity and reliability of a smartphone application for self-measurement of active shoulder range of motion in a standing position among healthy adults

Hiroki Shimizu, PT, MS, Takanobu Saito, PT, BS, Chika Kouno, PT, BS, Kanako Shimoura, PT, PhD, Rika Kawabe, MS, Yuki Shinohara, PT, MS, Kohei Mukaiyama, PT, MS, Chen Changyu, MS, Masaya Kato, OT, MS, Momoko Nagai-Tanima, PT, PhD, Tomoki Aoyama, MD, PhD

Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Keywords:
Shoulder range of motion
Goniometer
Smartphone
Intraclass correlation coefficient
Validity
Reliability

Level of evidence: Basic Science Study; Validation of Clinical Measurement Tool

Background: Shoulder range of motion (ROM) is one of the most important indicators of shoulder disease severity, function, and physical assessment. A universal goniometer (UG) was used as a gold standard for ROM measurement. Recently, smartphone applications for ROM measurement have attracted attention as alternatives to UG. This study aimed to investigate the validity and reliability of active ROM measurements using a smartphone application goniometer that can be used by patients in a standing position.

Methods: The dominant shoulders of 19 healthy participants were included in the study. The 2 observers who were physical therapists used the UG, whereas the participants used a smartphone application goniometer to measure the shoulder ROM. A recorder, who is a physical therapist independent of the observer and participant, read and recorded the shoulder ROM measurements. The order of the measurement movements and devices used was randomized.

Results: Agreement between the smartphone application goniometer and UG (percentage of participants for whom the difference between the UG and application measurements was within ±20% of the mean of the goniometer and application measurements) ranged between 42% and 100%. The intraclass correlation coefficient values (3, 1) for the agreement between the smartphone application goniometer and UG was between 0.72 and 0.97, showing significant and approximately perfect correlations.

Conclusion: High agreement with the UG showed excellent validity, indicating that the smartphone application goniometer used by the participants in the standing position is an excellent method and instrument. The results suggest a simpler, more reliable, practical, and inexpensive method for measuring ROM required for telerehabilitation.

© 2022 The Authors. Published by Elsevier Inc. on behalf of American Shoulder and Elbow Surgeons. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
reliability and validity of measuring shoulder ROM using a smartphone application and reported that the intraclass correlation coefficient (ICC) values for interobserver and intraobserver reliability were 0.80 and 0.89, respectively, showing a significant correlation. Additionally, Mejia-Hernandez et al reported an ICC value of >0.84 as the interobserver reliability of measuring shoulder ROM using a smartphone application. However, there are problems with these smartphone applications. The mobile device must be in contact with the patient and secured to the forearm using an armband. The armband needs to be disposable, as repeated use may cause hygiene issues. Additionally, the armband may loosen during use, requiring vigilance from the measurer. Therefore, it will be interesting to validate an evaluation method that does not require an item other than a smartphone, such as an armband. Additionally, the ROM measurements in previous studies were mainly performed in the supine position, which required time and space to lie down. To be faster and more suitable for daily life, it will be helpful to measure the ROM for all movements in a standing position. However, no studies have validated the accuracy of such measurements.

In recent years, research on telerehabilitation interventions has been increasing. Interventions are important in rehabilitation, and it is also important to evaluate physical functions before and after interventions to understand the patient's problems and measure the treatment effectiveness. Previous research on telerehabilitation assumes that only the intervention is performed remotely and that evaluation is conducted face to face at an actual facility. However, an active ROM assessment method that will allow patients to perform the entire process at home by themselves will be helpful to realize telerehabilitation. Many studies have been conducted to verify the accuracy of shoulder ROM measurements using the smartphone application goniometer; however, these are passive ROM measurements performed by observers, and no studies have verified the accuracy of active ROM measurements performed by participants themselves.

Therefore, this study aimed to evaluate the validity and reliability of active ROM measurement using a smartphone application goniometer that can be used by patients themselves in a standing position. We hypothesized that active shoulder ROM measurements using a smartphone application goniometer is as valid and reliable as ROM measurement using a UG.

Materials and methods

Participants

The participants were students recruited from the Kyoto University. The inclusion criteria were as follows: (1) age >20 years, (2) standing position during measurement, and (3) active movement of at least 90° of shoulder joint flexion and abduction. The exclusion criteria were as follows: (1) presence of pain in the glenohumeral, acromioclavicular, or sternoclavicular joint during active shoulder joint movement; and (ii) self-reported previous significant injury in the glenohumeral, acromioclavicular, or sternoclavicular joint. The dominant arm was used for writing, and the shoulder joint ROM of the dominant arm was measured.

Observers

Two physical therapists (H.S. and T.S.) who served as observers measured the shoulder ROM using a UG, and a recorder (C.K.), a physiotherapist independent of them, read and recorded the shoulder ROM to blind the results from the observers and participants. The week before the measurements were taken and on the day of the measurements, the 2 observers practiced the shoulder ROM measurements adequately.

Instruments

Active shoulder ROM was measured using a UG and smartphone application goniometer. The UG, Goniometer Todai-shiki 30 cm KO (Tsutsumi, Chiba, Japan) is a high-resolution stainless-steel goniometer that permits measurement of the axis of motion and joint ROM. The smartphone application goniometer, yROM (Healthcare Technologies LLC), is a smartphone application that sets the angle to 0° once the user touches the screen of the smartphone. The angle is measured when the user touches the screen again after moving. The smartphone application can be used by the participants themselves to measure the ROM.

Procedures

The following 8 active shoulder movements were measured: forward flexion (FF), extension (EXT), abduction (ABD), external rotation with the arm at the sides, external rotation at 90° ABD (ER90ABD), internal rotation at 90° ABD (IR90ABD), external rotation at 90° FF (ER90FF), and internal rotation at 90° FF (IR90FF). All the movement measurements were taken in a standing position, and the participants were instructed to place the back of their heads, backs, buttocks, and heels against a wall to prevent their trunk from moving. Each movement was measured 3 times by 2 observers on the goniometer and 3 times by the participant on the smartphone application. To eliminate the effects of learning and fatigue, the order of the measurement movements and equipment used was randomized. To control for the effects of poor movement and stretching, the participants practiced the movements thrice before each movement was measured. The participants were verbally instructed to hold their wrists in a neutral position to reduce the effect of the wrist joint. An IR90ABD was additionally measured under 2 conditions to find improvements in the measurement method using the smartphone application goniometer: First, the wrist was immobilized in an orthotic to verify the effect of wrist motion on measurement accuracy (IR90ABD [orthotic]); and second, an observer performed the application tapping motion required for the measurement on behalf of the participants to verify the effect of hand motion during tapping (IR90ABD [tap]).

Statistical analytics

Statistical analyses were performed using SPSS version 23 software (IBM Corp., Armonk, NY, USA). To validate the measurement of ROM using the smartphone application, the measurements obtained by the smartphone application were compared to those obtained by the goniometer. The smartphone application agreement relative to the goniometer was performed using the Brand-Altman analysis. When the number of participants for whom the difference between the goniometer and application measurements was within ±20% of the mean of the goniometer and application measurements was 75% or more of all the participants, the goniometer and application measurements can be considered to be in agreement. The validity of the smartphone application relative to the goniometer was calculated using the ICC (3, 1). The ICC values were calculated from the average of the 3 measurements of each observer using a goniometer and the average of the 3 measurements of each participant using the application. According to the guidelines of Landis and Koch, the ICC values of 0.00-0.20 are
considered to be slight correlations, 0.21–0.40 fair correlations, 0.41–0.60 moderate correlations, 0.61–0.80 substantial correlations, and 0.81–1.00 approximately perfect correlations.\(^{10}\)

Fixed and proportional biases of the smartphone application relative to the goniometer were evaluated using a Brand-Altman plot. The fixed bias was calculated using the 95% confidence interval (CI) of the difference between the UG and smartphone application goniometer measurements. If the 95% CI was zero, there was no fixed bias. If the 95% CI was <0, the smartphone application goniometer indicated underestimation; if the 95% CI was >0, the smartphone application goniometer indicated overestimation. The proportional bias was determined by the correlation coefficient calculated using the mean and difference between the smartphone application goniometer and UG.

The ICC (1, 1) was used to check the within-day reliability of the smartphone applications. The ICC values were calculated using 3 measurements for each participant.

To quantify the variability and measurement error, the standard error of measurement (SEM) and minimal detectable change (MDC) at the 95% CI were calculated. The SEM was calculated as SD/√2, and MDC\(_{95}\) was calculated as 1.96 × SD, where SD is the standard deviation of the difference between the 3 measurements of the application.\(^{6}\) Statistical significance was set at \(P < .05.\)

### Results

**Table I** presents the characteristics of the study participants. We analyzed the data from 19 shoulders (17 right and 2 left) of the 19 participants. The participants included 5 men and 14 women aged 22–42 years.

The results of the Bland-Altman analysis, where the percentage of participants for whom the difference between UG and application measurements was within ±20% of the mean of the goniometer and application measurements, are presented in **Table II**. The agreement between the smartphone application goniometer and UG ranged between 42% and 100%. The level of agreement for all movements was >75%, except for ER90ABD.

The ICC (3, 1) values for the agreement between the smartphone application goniometer and UG to verify the reliability of the smartphone application goniometer are reported in **Table III**. The ICC value for EXT was 0.72, indicating substantial correlations, whereas those for other movements ranged between 0.81 and 0.97, indicating approximately perfect correlations.

**Table IV** presents the 95% CIs for the differences between the UG and smartphone application goniometer measurements. The 95% CI for ABD was <0. Therefore, the smartphone application goniometer underestimated the ROM for ABD. The 95% CIs of ER90ABD by observer 1, IR90ABD by observers 1 and 2, and ER90FF by observer 2 were >0, and the smartphone application goniometer indicated overestimation in these ROM measurements.

The correlation coefficients of the difference between the UG and smartphone application goniometer measurements are presented in **Table V**. Significant correlations were found in EXT by observer 2, ABD by observer 1, and ER90FF by observer 1 and 2 with correlation coefficients of 0.46–0.66, indicating moderate correlation. Therefore, the smartphone application goniometer indicated a proportional bias in the ROM measurements.

The ICC (1, 1) values for the reliability of the 3 shoulder ROM measurements using the smartphone application goniometer are reported in **Table VI**. The ICC value for the EXT was 0.79, indicating substantive correlations, whereas those for the other movements were 0.89–0.95, indicating approximately perfect correlations.

### Table I

| Characteristics | Mean (SD) |
|-----------------|-----------|
| Age, yr         | 25.7 (4.4) |
| Body mass index, kg/m² | 21.2 (1.9) |

### Table II

| Movement | Observer 1 (%) | Observer 2 (%) |
|----------|----------------|----------------|
| FF       | 100            | 100            |
| EXT      | 79             | 100            |
| ABD      | 100            | 100            |
| ERS      | 100            | 95             |
| ER90ABD  | 100            | 100            |
| IR90ABD  | 53             | 42             |
| ER90FF   | 89             | 84             |
| IR90FF   | 100            | 100            |

**Table III**

| Movement | ICC (3, 1) (95% confidence interval) |
|----------|--------------------------------------|
| Observer 1 | Observer 2 |
| FF        | 0.94 (0.84–0.97) | 0.94 (0.84–0.96) |
| EXT       | 0.72 (0.40–0.89) | 0.78 (0.51–0.91) |
| ABD       | 0.94 (0.86–0.98) | 0.97 (0.92–0.99) |
| ERS       | 0.91 (0.78–0.96) | 0.85 (0.66–0.94) |
| ER90ABD   | 0.91 (0.77–0.96) | 0.94 (0.85–0.98) |
| IR90ABD   | 0.87 (0.70–0.95) | 0.85 (0.65–0.94) |
| ER90FF    | 0.83 (0.61–0.93) | 0.93 (0.84–0.97) |
| IR90FF    | 0.85 (0.64–0.94) | 0.81 (0.56–0.92) |

### Table IV

| Movement | 95% Confidence interval |
|----------|-------------------------|
| Observer 1 | Observer 2 |
| FF        | –0.72 to 5.10          | –1.36 to 4.13 |
| EXT       | –3.11 to 1.36          | –1.61 to 1.85 |
| ABD       | –7.33 to –2.07         | –5.68 to –1.65 |
| ERS       | –1.59 to 2.22          | –3.59 to 0.96 |
| ER90ABD   | 0.31 to 5.84           | –0.50 to 4.32 |
| IR90ABD   | 2.56 to 7.58           | 3.52 to 8.98  |
| ER90FF    | –0.39 to 0.48          | 0.40 to 3.38  |
| IR90FF    | –0.65 to 4.52          | –0.61 to 4.89 |

**Table V**

**Table VI**
The proportional biases of the smartphone application relative to the goniometer: a Brand-Altman plot.

| Movement                  | Correlation coefficients (P value) | Observer 1 | Observer 2 |
|---------------------------|-----------------------------------|------------|------------|
| FF                        | 0.13 (.59)                        | 0.42 (.07) |            |
| EXT                       | 0.04 (.88)                        | 0.56 (.01) |            |
| ABD                       | 0.53 (.02)                        | 0.36 (.13) |            |
| ERS                       | 0.13 (.60)                        | 0.30 (.21) |            |
| ER90ABD                   | 0.32 (.18)                        | -0.07 (.78)|            |
| IR90ABD                   | 0.40 (.09)                        | 0.37 (.12) |            |
| IR90FF                    | 0.46 (.046)                       | 0.24 (.34) |            |
| IR90FF                    | 0.47 (.04)                        | 0.66 (.002) |            |

*P value < .05.

The SEM for the 3 shoulder ROM measurements by the smartphone application goniometer ranged from 1.9 to 2.7, and the MDC90 ranged from 5.2 to 7.5 (Table VII).

Table VIII shows the results of the Brand-Altman analysis of additional measurements. The agreement of IR90ABD (orthotic) with the UG measurements by observers 1 and 2 was 62.5%, below the 75% level of agreement but above the IR90ABD agreement levels of 53% and 42%, respectively. The agreement between the IR90ABD (tap) measurement and goniometer measurements by observer 1 was 75%. The agreement of IR90ABD (tap) with the goniometer measurement by observer 2 was 62.5%, which was lower than the 75% level of agreement but higher than the IR90ABD agreement level of 42%.

Discussion

This study assessed the validity and reliability of a smartphone application goniometer, which participants themselves can use to measure ROM in a standing position. The agreement between the smartphone application goniometer and UG measurements in this study exceeded the 75% level of agreement for all the movements, except IR90ABD. The results indicate that FF, EXT, ABD, external rotation with the arm at the sides, ER90ABD, ER90FF, and IR90FF are validated measurements obtained using a smartphone application goniometer. The ICC (3, 1) values for the smartphone application goniometer compared with the UG ranged between 0.72 and 0.97. They were classified in the range of substantial to approximately perfect correlations. The ICC values ranged from 0.72 to 0.97, which is comparable to ICC values of 0.79-0.99 in a previous study examining the accuracy of smartphone-based passive goniometry in a supine position. The SEM ranged between 1.9 and 2.7, and MDC90 ranged between 5.2 and 7.5. The SEM and MDC90 showed lower values than those reported in previous studies. These results suggest that the active shoulder ROM measurements using a smartphone application goniometer in a standing position are as reliable as passive measurements in a supine position, which is currently the standard method. The ICC (1, 1) values for intraday reliability were 0.79-0.95 and classified as substantive correlations or approximately perfect correlations. These results indicate that the smartphone application goniometer had superior reproducibility.

This smartphone application goniometer has some excellent features; however, there are some limitations to its use. The results of the Brand-Altman plot showed that the smartphone application goniometer may have a fixed bias, which may underestimate ABD and overestimate ER90ABD, IR90ABD, and ER90FF. The smartphone application goniometer may have significant proportional biases in EXT, ABD, ER90FF, and IR90FF. The measurement of these movements may only be used in limited situations, such as pre-intervention and post-intervention evaluations.

Additional measurements were conducted on IR90ABD, which fell below the 75% agreement as a result of the Brand-Altman analysis. The agreement of IR90ABD (orthotic) with the UG measurements by observers 1 and 2 was 62.5%, above the IR90ABD agreement level of 53% and 42%, respectively. This result indicates that compensatory movement by the wrist joints may have occurred in the IR90ABD measurement, thereby reducing the agreement between the smartphone application goniometer and UG measurements. The agreement between the IR90ABD (tap) and goniometer measurements by observer 1 was 75%. The agreement of IR90ABD (tap) with the goniometer measurement by observer 2 was 62.5%, which was higher than the IR90ABD agreement level of 42%.
participant tapping the screen; however, a measurement method that automatically measures the joint angle by holding it for a certain period of time without requiring tapping may be better. Furthermore, this method of measurement may not necessarily be a concern because the results of this study showed similar level of reliability as those of the previous studies, although the smartphone application goniometer was fixed to the participant’s forearm while the examiner read the measurements in previous studies; thus, there was no effect of the participant’s wrist joint movement or tapping.

This study had several limitations. First, the participants were healthy individuals. It is assumed that the actual users of the application are symptomatic. Symptomatic patients may move differently from healthy people, and the accuracy of shoulder joint measurements may vary. Previous studies have confirmed that the accuracy of shoulder ROM measurement using smartphone applications is excellent even in symptomatic patients. However, the accuracy of standing and active shoulder ROM measurements in symptomatic patients has not been confirmed, and it requires further validation. This is because compensatory movements at the elbow and scapulothoracic joints may impact this assessment in injured shoulders. Therefore, the ROM measured in this study is limited to the fact that it is not a true scapulohumeral joint ROM but a functional ROM that includes the scapulohumeral joint. Second, there was a bias in the characteristics of the study participants. Most of the measurement sides were the right shoulder, and the shoulder joint movement may vary, depending on whether it is on the dominant or non-dominant side. These biases may have affected the measurement accuracy. Additionally, most of the participants were young and female. Patients with frozen shoulder who are assumed to be users of smartphone application goniometer are more likely to be female rather than male and are more common in the 40-60 years age group. It may be difficult to generalize the results for men and the elderly as most of the participants in this study were women, but it is not necessarily a major issue in frozen shoulder, which has a high incidence in female and nonelderly patients. Third, the sample size of this study was small. This study met the sample size required for IEC. However, a larger sample size allows for a more detailed analysis of the effects of the dominant arm, sex, and age.

Conclusion

This study evaluated the validity and reliability of using a smartphone application goniometer by the participants themselves to measure active shoulder ROM in a standing position. The results showed excellent validity, owing to high agreement with the UG and excellent intraday reliability values. This indicates that the smartphone application goniometer used by participants in a standing position is an excellent method and instrument. The results suggest a simpler, more reliable, practical, and inexpensive method for measuring ROM required for telerehabilitation.

Disclaimers:

Funding: No funding was disclosed by the authors.
Conflicts of interest: The authors, their immediate families, and any research foundation with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

Acknowledgments

The authors would like to express their sincere thanks to all participants who co-operated in the study, the students, and the laboratory members for their advice on submitting the study. They also thank Editage (www.editage.com) for the English language editing.

References

1. Albrecht UV, van Jon U, Sedlacek L, Groos S, Suerbaum S, Vonberg RP. Standardized, App-based disinfection of iPods in a clinical and nonclinical setting: comparative analysis. J Med Internet Res 2013;15:e176. https://doi.org/10.2196/jmir.2642.
2. Clarnette RG, Miniati A. Clinical exam of the shoulder. Med Sci Sports Excell 2008;30:51-6. https://doi.org/10.1097/00005768-199804010-00001.
3. Correll S, Field J, Hutchinson H, Mickeyvicius G, Fitzsimmons A, Smith B. Reliability and validity of the halo digital goniometer for shoulder range of motion in healthy subjects. Int J Phys Sports Ther 2018;13:707-14.
4. Critchley LA, Critchley JA. A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. J Clin Monit Comput 1999;15:85-91.
5. Fayad F, Roby-Brami A, Yazbeck C, Hamnetto S, Leefevre-Colau MM, Gautheron V, et al. Three-dimensional scapular kinematics and scapulohumeral rhythm in patients with glenohumeral osteoarthritis or frozen shoulder. J Biomech 2008;41:326-32. https://doi.org/10.1016/j.jbiomech.2007.09.004.
6. Haley SM, Fraga-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther 2008;86:735-41.
7. Han C, Clipsham K, Rees JL, Carr AJ. Long-term outcome of frozen shoulder. J Shoulder Elbow Surg 2008;17:231-6. https://doi.org/10.1016/j.jse.2007.05.009.
8. Hayes K, Walton JR, Szmor ZL, Murrell GA. Reliability of five methods for assessing shoulder range of motion. Aust J Physiother 2001;47:289-94.
9. Hosseinmehr SH, Abarani M, Norasteh AA, Fardmal J, Khorasveti MT. The comparison of scapular upward rotation and scapulohumeral rhythm between dominant and non-dominant shoulder in male overhead athletes and non-athletes. Man Ther 2015;20:758-62. https://doi.org/10.1016/j.math.2015.02.010.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
11. Lew R, Gherghe D. Design based on intra-class correlation coefficients. Am J Biostat 2010;1:1-8. https://doi.org/10.3844/amjbsp.2010.1.8.
12. Mejia-Hernandez K, Chang A, Eardley-Harris N, Jaarsma R, Gill TK, McLean JM. Smartphone applications for the evaluation of pathologic shoulder range of motion and shoulder scores—a comparative study. JSES Open Access 2018;2:102-14. https://doi.org/10.1016/j.jses.2017.10.001.
13. Mullaney MJ, McHugh MP, Johnson CP, Tyler TF. Reliability of shoulder range of motion comparing a goniometer to a digital level. Physiother Theor Pract 2010;26:327-33. https://doi.org/10.3109/09593980903094420.
14. Riddle DL, Rothstein JM, Lamb RL. Goniometric reliability in a clinical setting. Shoulder measurements. Phys Ther 1987;67:668-73.
15. Russell TG. Physical rehabilitation using telemedicine. J Telemed Telecare 2007;13:217-20. https://doi.org/10.1258/135761307781458886.
16. Shin SH, Ro DH, Lee OS, Oh JH, Kim SH. Within-day reliability of shoulder range of motion measurement with a smartphone. Man Ther 2012;17:298-304. https://doi.org/10.1016/j.math.2012.02.010.
17. Tekerekolu MS, Duman Y, Serindag A, Caglar SS, Kayasud H, Tunc E, et al. Do mobile phones of patients, companions and visitors carry multidrug-resistant hospital pathogens? Am J Infect Control 2011;39:379-81. https://doi.org/10.1016/j.ajic.2010.10.026.
18. Werner BC, Holzgreve RE, Griffin JW, Lyons M, Cosgrove CT, Hart JM, et al. Validation of an innovative method of shoulder range-of-motion measurement using a smartphone clinometer application. J Shoulder Elbow Surg 2014;23:1275-82. https://doi.org/10.1016/j.jse.2014.02.030.
19. Yes SM, Lim JY, Do JG, Lim J, Lee J, Iwung JH. Effectiveness of interactive augmented reality-based telerehabilitation in patients with adhesive capsu- litis: protocol for a multi-center randomized controlled trial. BMC Musculoskelet Disord 2021;22:386. https://doi.org/10.1186/s12891-021-04261-1.