



Original Article

## Accuracy and reliability of a smartphone application for measuring the knee joint angle

KEISUKE ISHII, MD<sup>1)\*</sup>, HIROYUKI OKA, MD, PhD<sup>2)</sup>, YUJI HONDA, MD<sup>3)</sup>, DAISUKE OGURO, MD<sup>3)</sup>, YOUICHIRO KONNO<sup>4)</sup>, KOUSUKE KUMETA, BPhy<sup>4)</sup>, SHOUTA NISHIHARA<sup>4)</sup>, HAZUKI MATSUYAMA, BPhy<sup>4)</sup>, ICHIRO KANEKO, MD, PhD<sup>5)</sup>, YASUO TAKEUCHI, MS<sup>6)</sup>, YOSHINOBU WATANABE, MD, PhD<sup>1)</sup>, HIROTAKA KAWANO, MD, PhD<sup>7)</sup>, NAOSHI OGATA, MD, PhD<sup>3)</sup>

<sup>1)</sup> Trauma and Reconstruction Center, Teikyo University Hospital: 2-11-1 Kaga, Itabashi-Ku, Tokyo 173-8606, Japan

<sup>2)</sup> 22nd Century Medical and Research Center, The University of Tokyo Hospital, Japan

<sup>3)</sup> Department of Rehabilitation Medicine, Teikyo University School of Medicine, Japan

<sup>4)</sup> Rehabilitation Unit, Teikyo University Hospital, Japan

<sup>5)</sup> Department of Emergency Medicine, Teikyo University School of Medicine, Japan

<sup>6)</sup> Teikyo Simulation Education Research Center, Teikyo University School of Medicine, Japan

<sup>7)</sup> Department of Orthopaedic Surgery, Teikyo University School of Medicine, Japan

**Abstract.** [Purpose] Recently, a photo-based smartphone application for angle measurement—“Grid line imaging application Professional”—was developed to evaluate joint disease treatments. The aim of this study was to determine the accuracy and reliability of the application. [Participants and Methods] We measured the knee joint of a mannequin using an application and a universal goniometer. Twelve examiners measured eight knee joints of mannequins at different arbitrary angles using the application and a universal goniometer. Correlations between the application and universal goniometer measurements were examined using scatter plots and correlation coefficients. Systematic errors of the application were visually confirmed using the Bland-Altman method. Intra-class correlation coefficients were used to evaluate the inter-examiner reliability of the application. [Results] The application and universal goniometer measurements showed a good correlation ( $r=0.99$ ) and no systematic error. The intra-class correlation coefficient for inter-examiner reliability was 0.999. Furthermore, to evaluate intra-examiner reliability, six examiners measured six different knee joints twice using the application on a 2-day interval. The intra-class correlation coefficient for intra-examiner reliability was 0.982. [Conclusion] The accuracy of the application was equivalent to that of a universal goniometer, and both the inter- and intra-examiner reliabilities of the application were almost perfect.

**Key words:** Angle measurement, Smartphone application, Reliability

(This article was submitted Jan. 7, 2021, and was accepted Feb. 17, 2021)

### INTRODUCTION

Range of motion measurements are widely used in evaluating patients with bone, joint, muscle, and neurological disorders. The Japanese Orthopaedic Association and the Japanese Society of Rehabilitation Medicine have established a method for measuring range of motion using the angle between the basic axis and the axis of motion of the joint in one plane, using the neutral starting position (0 degrees) as the basic position<sup>1)</sup>. For the knee joint, the basic limb position is set at 0 degrees of extension in the sagittal plane, the basic axis is the femur, and the axis of translation is the fibula. The range of motion

\*Corresponding author. Keisuke Ishii (E-mail: keisuke99ishii@gmail.com)

©2021 The Society of Physical Therapy Science. Published by IPEC Inc.



This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

is assessed in terms of maximum extension and maximum flexion angles. During treatment of musculoskeletal trauma as well as neurological or joint diseases, joint angle of motion is frequently measured over time to assess the effectiveness of the treatment. Range of motion is also used to formulate a diagnosis for insurance and welfare purposes. Thus, joint angle measurement is an important routine procedure for physicians as well as physical and occupational therapists. Usually, a universal goniometer (UG) is used to measure joint angles.

The use of a UG was evaluated for good reliability and validity in the 1980s<sup>2-5</sup>). For instance, the 300-mm UG is used to accurately capture the proximal and distal landmarks of the basic axis for angular measurements of joints with long basic axes, such as the knee joint; however, it is inconvenient to carry around due to its size.

With the development of technology, more people use their smartphones to record measurements since they are easy to carry around. In recent years, various smartphone applications for angular measurement have been developed. These include accelerometers<sup>6, 7</sup>), gyroscopes<sup>8</sup>), magnetometers<sup>9</sup>), and photography (photo-based type)<sup>10</sup>) that measure joint angles and they have all been examined for validity and reliability. In 2017, the Japanese-language smart phone application “Grid Line Photography App Professional” (APP, Naradewa Inc) was developed for angular measurement. It is used to measure the angle between two straight lines by taking a photo with a smart phone and determining the three arbitrary points on the photo. It is easy to use due to its familiarity because the process of taking a photograph close to a joint and determining its three points is similar to the widely used process of measuring joint angles using a UG. However, the measurements resulting from the APP will only be accurate if the photograph is taken perpendicularly to the plane containing the axis of movement of the joint<sup>11</sup>). In addition, if the three required points for joint angle measurement are not accurately determined on the photograph, the measurement will be inaccurate.

The validity and reliability of photo-based joint angle measurements have already been evaluated<sup>10, 11</sup>). However, the accuracy and reliability of the APP has not been established. Thus, the purpose of this study was to determine the accuracy and reliability of the APP.

## PARTICIPANTS AND METHODS

We measured the knee joint angle of a mannequin (MAN; Resusci Anne Simulator, model number 150-20049) used for emergency resuscitation training using the APP and a UG (MMI universal goniometer Todai 300 mm, Muranaka Medical Instruments, Co., Ltd., 18/8 stainless steel). The APP was used to measure the knee joint angle using a smartphone (iPhone 8, Apple, iOS13.1, Model: A1906, Size: 138.4 mm long, 67.3 mm wide, 7.3 mm thick); it was downloaded on a smartphone, which was then used to take pictures of the mannequin’s knee joint from the lateral side. The points assumed to be the center of the knee joint (Fig. 1, point B) and the external fruit of the ankle joint (Fig. 1, point C) were specified. Meanwhile, the knee joint angle was measured using the UG by placing its center at the center of the knee joint of the mannequin with one axis of the UG placed on the line considered to be the femoral axis and the other placed on the line considered to be the lower leg axis.

The eight knee joints of four mannequins were fixed at different arbitrary angles to prevent them from moving. The angles were set to be unbiased, ranging from 0 degrees to the maximum flexion angle of the mannequin knee joint (100 degrees). All eight knee joint angles were each measured by 12 examiners using the APP. The examiners included nine physical therapists, two occupational therapists, and one physician. They routinely used a UG in their practice and had never used the APP; the instructions regarding the use of the APP were given to all of them just before the commencement of measurements. Thereafter, a UG was used to measure the same eight knee joints, which were recorded at 15 minutes intervals. Each examiner could not know the measurements of the other examiners.

Thus, the 12 examiners measured the eight knee joints using both the APP and a UG, resulting in 96 pairs of paired data. Correlations between the APP and UG measurements were examined by obtaining scatterplots and correlation coefficients. Accuracy was examined by visually confirming the systematic errors of the APP using the Bland-Altman method<sup>12</sup>). We also evaluated the inter-examiner reproducibility of the APP and the UG by calculating the intra-class correlation coefficients (ICC).



**Fig. 1.** Measurement of knee joint angle using a smartphone application.

Six knee joints of three mannequins were set at different arbitrary joint angles and firmly placed on a platform so that the angles would not change. These joints were then measured and recorded by six examiners using the APP. A second measurement was performed after three days by the same examiners. The examiners included four physiotherapists, one occupational therapist, and one physician. All six examiners were included in the same list of examiners who assessed the correlation of the APP and UG. Six pairs of data were obtained from each examiner.

Accuracy was judged based on the correlation coefficients between the APP and the UG measurements, standard deviation of the Bland-Altman plot, and 95% confidence intervals (CIs). The inter- and intra-examiner reproducibility was examined using ICC. The ICCs were classified as follows: almost perfect (0.81–1.0), substantial (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40), and slight (0–0.20)<sup>13</sup>. Statistical analyses were performed using IBM SPSS Statistics Version 24.0, IBM software (Armonk, NY, USA: IBM Corp). The significance level was set at p-value <0.05.

## RESULTS

The average and standard deviation of joint angle measurements of the eight knee joints, as measured by the 12 examiners using the APP and UG, are shown in Tables 1 and 2. Meanwhile, the scatterplot of 96 pairs of data from the APP and UG measurements, which are exhibited on the y- and x-axis, respectively, is shown in Fig. 2. The UG and APP measurements showed a good correlation (correlation coefficient  $r=0.99$ , 95% CI: 0.98–0.99, p-value<0.001).

When the systematic errors of the APP and UG were visualized using the Bland-Altman method, there was no bias either above or below the standard line (Fig. 3). The good correlation between the APP and UG as well as the absence of phylogenetic errors indicate a high accuracy of the APP.

We evaluated the inter-examiner reproducibility of the APP and UG by determining the ICCs (95% CIs), which were as follows: for single measurements, APP=0.987 (0.976–0.998) and UG=0.986 (0.966–0.997) (Table 3); for the average measurements, APP=0.999 (0.998–1.000) and UG=0.999 (0.997–1.000) (Table 4). The evaluations of the ICC were almost perfect.

**Table 1.** The knee joint angle measurements of APP in inter-examiner reproducibility tests

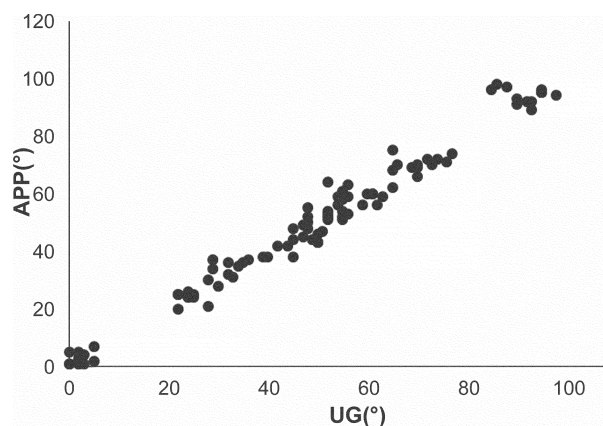
	MAN1	MAN2	MAN3	MAN4	MAN5	MAN6	MAN7	MAN8
Average	2.8	24.5	36.1	46.4	54.3	58.3	70.5	94.0
SD	1.9	2.5	2.8	3.5	4.6	3.1	2.4	2.5

APP: smartphone application; MAN: knee joint angle of mannequin; SD: standard deviation.

**Table 2.** The knee joint angle measurements of UG in inter-examiner reproducibility tests

	MAN1	MAN2	MAN3	MAN4	MAN5	MAN6	MAN7	MAN8
Average	2.2	24.8	35.5	48.0	52.5	58.9	70.6	91.7
SD	1.6	2.4	4.7	2.2	2.4	3.3	3.7	3.6

UG: universal goniometer; MAN: knee joint angle of mannequin; SD: standard deviation.

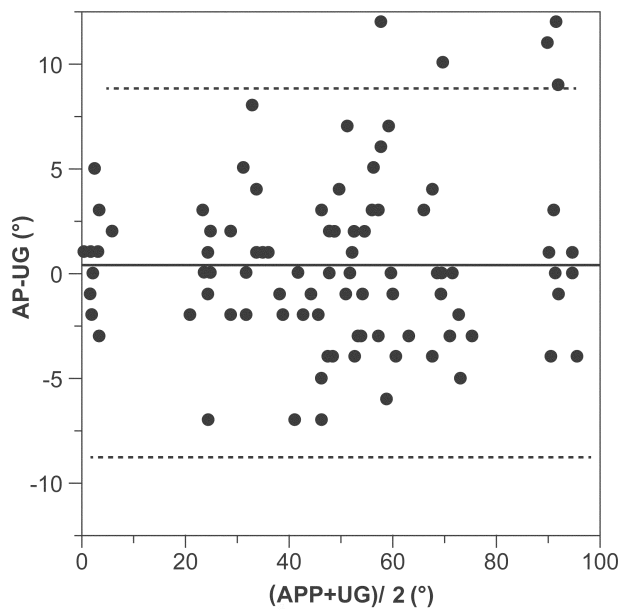


**Fig. 2.** Scatter plot of identical knee angle measurements by the APP and the UG. APP: smartphone application; UG: universal goniometer.

The measurements of the six knee joints, as measured by the six examiners using the APP, are shown in Table 5. The intra-examiner reproducibility of the APP was assessed using the ICC (95% CI), which was 0.982 (0.965–0.991) for single measurements and 0.991(0.982–0.995) for the average measurement (Table 6). The evaluation of the ICC was almost perfect.

## DISCUSSION

Several studies have evaluated smartphone applications that measure joint angles<sup>14</sup>). However, no evaluation of a photo-based application in Japanese exists. Thus, this study compared the use of the APP with a UG and demonstrated the accuracy of the APP, since they were well correlated and free from systematic errors. The reliability of the APP was also demonstrated by an almost perfect ICC both its inter- and intra-rater reproducibility.



**Fig. 3.** Bland-Altman plot of identical knee angle measurements by the UG and the APP.

APP: smartphone application; UG: universal goniometer.

**Table 3.** ICC of inter-examiner reliability (single measure value)

	ICC	95% CI	p value
APP	0.987	0.976 to 0.998	<0.001
UG	0.986	0.966 to 0.997	<0.001

APP: smartphone application; UG: universal goniometer; ICC: intra-class correlation coefficients; CI: confidence interval.

**Table 4.** ICC of inter-examiner reliability (average measured value)

	ICC	95% CI	p value
APP	0.999	0.998 to 1.000	<0.001
UG	0.999	0.997 to 1.000	<0.001

APP: smartphone application; UG: universal goniometer; ICC: intra-class correlation coefficients; CI: confidence interval.

**Table 5.** The knee joint angle measurements of APP in intra-examiner reproducibility tests

	Tester 1		Tester 2		Tester 3		Tester 4		Tester 5		Tester 6	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
MAN1	13	13	10	9	13	17	14	20	15	10	10	10
MAN2	45	48	44	46	48	54	49	54	50	52	50	45
MAN3	64	62	60	58	68	65	65	67	65	60	60	60
MAN4	38	38	33	39	35	40	45	44	38	43	40	35
MAN5	53	53	52	53	57	56	60	59	54	52	55	55
MAN6	18	15	19	17	17	24	22	18	14	12	14	20

APP: smartphone application; MAN: knee joint angle of mannequin.

**Table 6.** ICC of intra-examiner reliability of APP

	ICC	95% CI	p value
Single measure value	0.982	0.965 to 0.991	<0.001
Average measured value	0.991	0.982 to 0.995	<0.001

APP: smartphone application; ICC: intra-class correlation coefficients; CI: confidence interval.

For the systematic errors of APP and UG, this study used the Bland-Altman method, as reported in several studies<sup>6, 8, 10, 15–18</sup>. There are two types of errors: systematic and chance errors. Systematic errors have a certain bias toward the true value, while chance errors occur randomly in either direction or magnitude in relation to the true value. Systematic errors become biased even after repeated measurements, which may lead to misinterpretation of the results. Thus, to address the systematic errors, the Bland-Altman plot was created; in the study, the existence and degree of systematic errors were visually confirmed, and the distribution showed no bias in either direction, indicating no systematic error between the APP and the UG.

The inter-examiner reproducibility of the eight knee joints that were measured by the 12 examiners using the APP had a high ICC of 0.999. Similarly, the intra-examiner reproducibility of the six knee joints measured twice by the six examiners using the APP at 2-day intervals was also high with an ICC of 0.982. An ICC of 0.81 or higher is considered almost perfect, indicating that the angular measurement taken with the APP were reliable for all examiners at one measurement.

In this study, we used the knee joint of a mannequins rather than a human participant to measure angles to avoid the ethical issues of using human participant. In addition, the angle to be measured was fixed and did not change during the measurements. The use of mannequins rather than humans prevented the knee joint angle from changing during the measurement by multiple examiners at two-day intervals. We chose the knee specifically because it is a hinge joint, making it easy to measure using both the APP and the UG. The hip, knee, and ankle joints of the mannequins used in this study could also be moved, and the range of motion of the knee joint ranged from 0 to 100 degrees. Therefore, the angles were set at various angles ranging from 0 to 100 degrees for both the inter-examiner (8 joints) and the intra-examiner (6 joints) reproducibility tests; however, the deep flexion angle could not be set above 100 degrees. Hence, whether the APP could measure knee joint deep flexion angle precisely was not shown. In addition, because the structure of the mannequin's lower extremity did not have an axis corresponding to either the femur or tibia, the basic axis and the axis of movement were measured based on appearance when measuring the knee joint angle. The mannequins were not naked but were clothed with shorts, as in an actual measurement of a patient's knee joint angle (Fig. 1).

Previous studies examining the intra-examiner reliability of joint angle measurements have often used multiple measurements by the same examiner on the same day<sup>6, 7, 15</sup>. However, in our study, the second measurement was performed three days later to prevent the examiner from remembering the first measurement angle, which could affect the second measurement. The knee joint angle of the MAN was strictly secured so that it would not change before the second measurement, and the room where it was placed was closed off between the first and second measurements. In all knee joint angle measurements, the angles measured by the examiners were recorded by and for the examiners themselves, which were then collected immediately after the measurement to prevent examiners from comparing their values. We believe that these considerations reduced the information bias.

There are several reports on the validity and reliability of smartphone applications for joint angle measurement; 37 articles were included in a systematic review by Keogh et al.<sup>14</sup> Nine of these papers obtained knee joint angle measurements<sup>6, 15, 17–23</sup> using six applications that were used as accelerometers, which were all evaluated for validity and reliability. Of these applications, two were used on a Samsung Galaxy<sup>21, 22</sup> device, while the other four were used on an Apple iPhone<sup>6, 15, 17–20, 23</sup>. However, unlike the accelerometer app, the photo-based app is intuitive and easy to use as it determines the angle of the basic axis as well as the measurement using a UG. In addition, it can measure angles without having contact with the patient, which is not only desirable in preventing infection but for remote diagnosis as well. It has been confirmed in the laboratory that angles can be accurately measured using photographs taken with smartphones<sup>24</sup>. However, angular measurements of a photo-based smartphone application might be inaccurate if the photo was taken in the wrong orientation<sup>10</sup>. Nevertheless, previous studies have shown that the reliability of photo-based type applications was high, with ICCs of 0.9 or higher both between and within the examiners<sup>10</sup>. In this study, we showed that the reliability of the APP is almost as good as that of a UG; however, to improve its accuracy and reliability, we are currently working on developing a system that uses artificial intelligence to automatically determine the three points of knee angle measurement and automate the angular measurement.

The limitation of this study is that we used a mannequin instead of a human participant for angular measurements. In the future, it is necessary to examine whether it is possible to have the same accuracy and reliability when performing angular measurements in humans. In addition, since this study was conducted only on knee joints, which are hinged joints, further studies are needed to determine whether the APP can be used as well on ball joints, such as the shoulder and hip joints.

This study demonstrated that the APP is as accurate and reliable as a UG in knee joint angle measurements.

### *Funding and Conflicts of interest*

Keisuke Ishii receives grants (No. 18097, 500,000 JPY) from Naradewa Inc. The remaining authors declare no conflict of interest.

## **ACKNOWLEDGMENT**

We would like to thank Editage ([www.editage.com](http://www.editage.com)) for the English language editing.

## REFERENCES

- 1) Yonemoto K, Ishigami S, Kondo T: Range of motion display and measurement methods. *Rihabiriteshon Igaku*, 1995, 32: 207–217 (In Japanese). [[CrossRef](#)]
- 2) Rothstein JM, Miller PJ, Roettger RF: Goniometric reliability in a clinical setting. Elbow and knee measurements. *Phys Ther*, 1983, 63: 1611–1615. [[Medline](#)] [[CrossRef](#)]
- 3) Gogia PP, Braatz JH, Rose SJ, et al.: Reliability and validity of goniometric measurements at the knee. *Phys Ther*, 1987, 67: 192–195. [[Medline](#)] [[CrossRef](#)]
- 4) Rheault W, Miller M, Nothnagel P, et al.: Intertester reliability and concurrent validity of fluid-based and universal goniometers for active knee flexion. *Phys Ther*, 1988, 68: 1676–1678. [[Medline](#)] [[CrossRef](#)]
- 5) Enwemeka CS: Radiographic verification of knee goniometry. *Scand J Rehabil Med*, 1986, 18: 47–49. [[Medline](#)]
- 6) Ockendon M, Gilbert RE: Validation of a novel smartphone accelerometer-based knee goniometer. *J Knee Surg*, 2012, 25: 341–345. [[Medline](#)] [[CrossRef](#)]
- 7) Keijsers R, Zwerus EL, van Lith DR, et al.: Validity and reliability of elbow range of motion measurements using digital photographs, movies, and a goniometry smartphone application. *J Sports Med (Hindawi Publ Corp)*, 2018, 2018: 7906875. [[Medline](#)]
- 8) Santos C, Pauchard N, Guilloteau A: Reliability assessment of measuring active wrist pronation and supination range of motion with a smartphone. *Hand Surg Rehabil*, 2017, 36: 338–345. [[Medline](#)] [[CrossRef](#)]
- 9) Johnson LB, Sumner S, Duong T, et al.: Validity and reliability of smartphone magnetometer-based goniometer evaluation of shoulder abduction—a pilot study. *Man Ther*, 2015, 20: 777–782. [[Medline](#)] [[CrossRef](#)]
- 10) Ferriero G, Vercelli S, Sartorio F, et al.: Reliability of a smartphone-based goniometer for knee joint goniometry. *Int J Rehabil Res*, 2013, 36: 146–151. [[Medline](#)] [[CrossRef](#)]
- 11) Dunlevy C, Cooney M, Gormely J: Procedural considerations for photographic-based joint angle measurements. *Physiother Res Int*, 2005, 10: 190–200. [[Medline](#)] [[CrossRef](#)]
- 12) Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1986, 1: 307–310. [[Medline](#)] [[CrossRef](#)]
- 13) Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics*, 1977, 33: 159–174. [[Medline](#)] [[CrossRef](#)]
- 14) Keogh JW, 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: e0215806. [[Medline](#)] [[CrossRef](#)]
- 15) Milanese S, Gordon S, Buettner P, et al.: Reliability and concurrent validity of knee angle measurement: smart phone app versus universal goniometer used by experienced and novice clinicians. *Man Ther*, 2014, 19: 569–574. [[Medline](#)] [[CrossRef](#)]
- 16) Modest J, Clair B, DeMasi R, et al.: Self-measured wrist range of motion by wrist-injured and wrist-healthy study participants using a built-in iPhone feature as compared with a universal goniometer. *J Hand Ther*, 2019, 32: 507–514. [[Medline](#)] [[CrossRef](#)]
- 17) Mehta SP, Barker K, Bowman B, et al.: Reliability, concurrent validity, and minimal detectable change for iPhone goniometer app in assessing knee range of motion. *J Knee Surg*, 2017, 30: 577–584. [[Medline](#)] [[CrossRef](#)]
- 18) Jones A, Sealey R, Crowe M, et al.: Concurrent validity and reliability of the simple goniometer iPhone app compared with the universal goniometer. *Physiother Theory Pract*, 2014, 30: 512–516. [[Medline](#)] [[CrossRef](#)]
- 19) Pereira LC, Rwakabayiza S, Lécurveux E, et al.: Reliability of the knee smartphone-application goniometer in the acute orthopedic setting. *J Knee Surg*, 2017, 30: 223–230. [[Medline](#)]
- 20) Hancock GE, Hepworth T, Wembridge K: Accuracy and reliability of knee goniometry methods. *J Exp Orthop*, 2018, 5: 46. [[Medline](#)] [[CrossRef](#)]
- 21) Dos Santos RA, Derhon V, Brandalize M, et al.: Evaluation of knee range of motion: correlation between measurements using a universal goniometer and a smartphone goniometric application. *J Bodyw Mov Ther*, 2017, 21: 699–703. [[Medline](#)] [[CrossRef](#)]
- 22) Derhon V, Santos RA, Brandalize M, et al.: Intra- and inter-examiner reliability in angular measurements of the knee with a smartphone application. *Human Mov*, 2017, 18: 38–43.
- 23) Hambly K, Sibley R, Ockendon M: 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. *Int J Physiother Phys Rehabil*, 2012, 2: 1–14.
- 24) Awatani T, Enoki T, Morikita I: Reliability and validity of angle measurements using radiograph and smartphone applications: experimental research on protractor. *J Phys Ther Sci*, 2017, 29: 1869–1873. [[Medline](#)] [[CrossRef](#)]