HUMAN MOVEMENT
2017; 18(2): 38–43

INTRA- AND INTER-EXAMINER RELIABILITY IN ANGULAR MEASUREMENTS OF THE KNEE WITH A SMARTPHONE APPLICATION

original paper
doi: 10.1515/humo-2017-0011

VIVIANE DERHON[1], RAFAEL APARECIDO SANTOS[1], MICHELLE BRANDALIZE[2], DANIELLE BRANDALIZE[1], LUCIANO PAVAN ROSSI[1]

[1] Central West State University, Guarapuava, Brazil
[2] Faculty Campo Real, Guarapuava, Brazil

ABSTRACT
Purpose. Goniometric smartphone applications to measure joint angles offer greater practicality and accessibility, which makes them potential alternatives to a conventional goniometer. The aim of the present study was to evaluate the intra- and inter-examiner reliability in measuring the angles of the range of motion of the knee with the use of the ROM© goniometric smartphone application.
Method. The total of 34 young healthy women with an at least 20° limitation in knee extension range of motion participated in the study. Angular measurements of knee flexion in the dominant leg were performed with the aid of the smartphone application by three trained examiners for the evaluation of the inter-examiner reliability. The second evaluation was carried out after a 48-hour period for the determination of the intra-examiner reliability.
Results. The proposed method demonstrated excellent intra-examiner (intra-class correlation coefficient [ICC] > 0.80) and inter-examiner (ICC > 0.90) reliability, with good intra-examiner ($r > 0.85$) and strong inter-examiner ($r > 0.90$) correlations ($p < 0.0001$), an intra-examiner coefficient of determination (R^2) of 0.75, inter-examiner R^2 of 0.91, and good level of agreement on the Bland-Altman plots for the measures.
Conclusions. The findings reveal that the ROM© goniometric smartphone application can be considered a useful tool for the evaluation of the knee range of motion in healthy women.
Key words: smartphone, joint range of motion, knee, reproducibility of measurements

Introduction

Thousands of applications in the fields of health and fitness are currently available on the market, including goniometric applications for measuring the range of joint motion [1]. Such applications ensure greater practicality and accessibility during physical evaluation [2] in a wide range of fields and offer advantages in terms of usability, with fewer technical difficulties [3].

Goniometric applications are based on the use of an accelerometer, gyroscope and magnetometer to transform the motion of body segments into numeric values, providing instantaneous readings [2]. Therefore, this new resource can become a fundamental tool, since the measure of range of motion is an important variable in physiotherapeutic evaluations for the establishment of joint functionality [4].

Owing to the importance of measures of joint movements for decision making during physiotherapeutic treatment, the reliability of the tools employed is an important aspect of precise evaluation [5]. Thus, such tools need to be submitted to thorough testing prior to their usage in clinical practice [6]. The use of novel clinical evaluation methods depends on the determination of the degree of inter- and intra-examiner reliability, as well as the degree of specificity in the measures of the equipment to be used [7]. The measurements should be performed in a standardized manner to enable comparisons during treatment and the presentation of the results to other health professionals [8].

Studies involving goniometric smartphone applications indicate a high degree of reliability for the knee in patients submitted to joint reconstruction surgery [3], in the comparison of experienced and novice examiners [6] and among inexperienced examiners [9]. Good to excellent reliability has also been found for the measurement of the range of motion of other joints with the use of smartphone applications [10–16]. While studies

Correspondence address: Luciano Pavan Rossi, Central West State University, St. Simeão Varela de Sá, 03, Vila Carli, 85040-080 Guarapuava, Paraná, Brazil, e-mail: lucianofisioo@yahoo.com.br

Received: February 17, 2017
Acepted for publication: April 7, 2017

Citation: Derhon V, Santos RA, Brandalize M, Brandalize D, Rossi LP. Intra- and inter-examiner reliability in angular measurements of the knee with a smartphone application. Hum Mov. 2017;18(2):38–43; doi: 110.1515/humo-2017-0011.

have demonstrated that smartphone applications are reliable for use in research and clinical practice [3, 6, 9–16], there is a wide diversity of goniometric applications to be tested. According to Milani et al. [1], further studies are needed for the validation of these applications for different joints and conditions.

The aim of the present study was to evaluate the intra- and inter-examiner reliability regarding the use of the ROM© smartphone goniometric application for angular measures of the knee in healthy women, since no study on this application has previously been published.

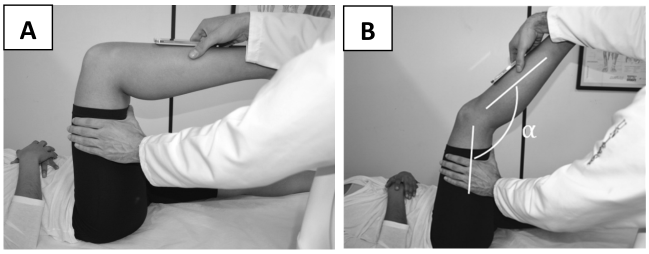


Figure 1. Determination of the knee angle with a smartphone application. A. Initial position (0°). B. Final position. α – the final joint angle of the knee flexion [17]

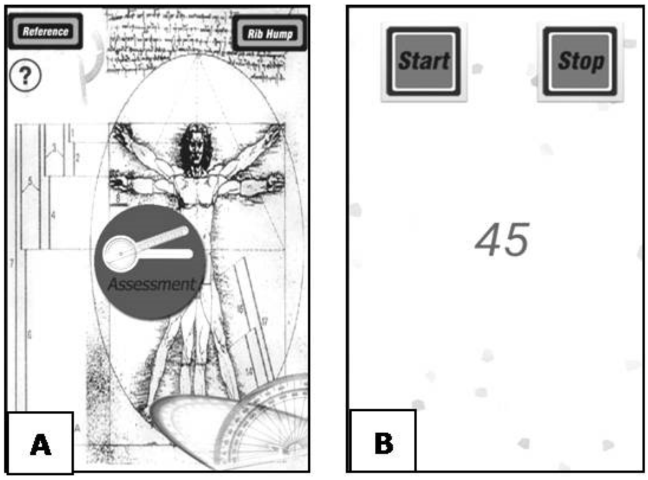


Figure 2. The ROM© application for the Samsung Galaxy smartphone. A. Initial screen. B. Screen for the joint angle measurement [17]

Material and methods

Participants

An analytical, correlational, blind, cross-sectional study was conducted with a sample of 34 young, healthy females (mean age: 21 ± 2 years; mean height: 163 ± 5.5 cm; mean weight: 60.2 ± 10.1 kg; and mean body mass index: 22.7 ± 3.5 kg/m^2). The inclusion criterion was a minimum limitation of 20° for the ROM of final knee extension of the dominant limb, measured with a universal goniometer, with the volunteer in the supine position and the hip flexed at 90° [17]. Individuals with pain or any lower limb disorder that could affect the evaluation of knee ROM were excluded from the study. The data collection was part of another study of the same authors [17] concerning a smartphone application for range of motion measurement. This study received an approval from the ethics committee of the Universidade Estadual do Centro-Oeste, Brazil (approval number: 1.015.221).

Procedures and tools

The measurements were performed by three trained examiners on two different days with a 48-hour interval, at the Physical Therapy Teaching Clinic of the Universidade Estadual do Centro-Oeste, Brazil. All examiners underwent a process of familiarization one week prior to the data collection process. The participants were also acquainted with the test before each evaluation, in the leg contralateral to the one to be evaluated.

At the measurement of the angle of knee range of motion, the participant remained in the supine position, with the hip and knee of the dominant limb flexed at 90° (initial position: 0°), with the foot relaxed, and the contralateral leg extended (Figure 1A). The knee was then passively extended by the examiner until the participant reported the onset of tension in the posterior muscle of the thigh (final position) (Figure 1B); this allowed to obtain the measures of the degree of knee flexors shortening.

The free ROM© goniometric application (v.1.4) for the Samsung Galaxy S5 smartphone (Figure 2A) was employed. The smartphone was positioned on the middle third of the anterior region of the shin [6] and held in place by the examiner during the movement from the initial position to the final position (Figure 1). In the initial position, the examiner touched the smartphone screen to determine the 0° position. Then they moved the leg and touched the screen again in the final position to record the angle produced by the movement (Figure 2B).

The order of the evaluations performed by different examiners was determined randomly with the Random Number Generator Pro 2.15© in a blinded fashion (the examiners were unaware of the other examiners' measurements). Each examiner carried out two evaluations, with a 30-second interval between measurements, and the mean of the two measurements was used for the analysis. In the pilot study, it was determined that more than two evaluations by three different examiners produced an increase in the flexibility of the posterior muscles of the thigh, which could exert an influence on the final results of the study.

Statistical analysis

The Shapiro-Wilk test and Levene's test were used to determine the normality and equality of variances of the data, respectively. On the basis of these tests results, the parametric statistical method was employed.

The comparison between the assessment tools was performed with the use of the paired t-test, with the level of significance set at 5% ($p < 0.050$).

The inter- and intra-examiner reliability was determined on the basis of the calculation of the intra-class correlation coefficient (ICC), with a 5% level of significance ($p < 0.050$), and the calculation of the 95% confidence intervals. The interpretation of the ICC was based on the Fleiss et al. pattern [18]: < 0.4 = low reliability; 0.4–0.75 = moderate to good reliability; and > 0.75 = excellent reliability. Pearson's correlation coefficients (r) were also calculated and interpreted in accordance with Rothstein et al. [19]: 0.90–0.99 = high correlation; 0.80–0.89 = good correlation; 0.70–0.79 = small correlation; and ≤ 0.69 = low correlation.

The coefficient of determination (R^2) was applied as a measure of the proportion of variability in one variable that is explained by the variability in another variable, demonstrating the predictive power of the associated regression line. The R^2 values range from 0 to 1, with values closer to 1 indicating that the proposed model is adequate for describing a given phenomenon [20].

Bland-Altman [21] agreement analysis was also employed. Measures were considered to be in agreement when the bias (difference between measures) was less than 5°, which is regarded as the minimum measurable difference that can have a clinically relevant impact [22], with the p-value > 0.050 and when 95% of the measures of the subjects were within the upper and lower limits of agreement [23].

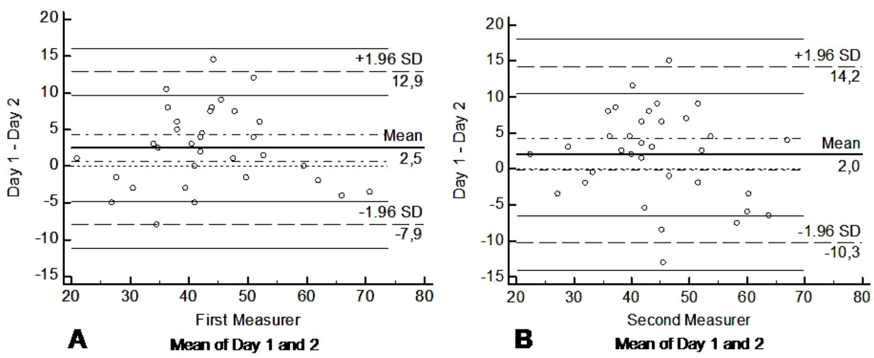
Results

For the intra-examiner data (Table 1), the paired t-test revealed a significant difference only for the first examiner, who obtained a mean difference of 2.5° between the two days. The intra-examiner reliability was considered excellent (ICC > 0.80) for all examiners ($p < 0.0001$).

Table 1. Intra-examiner analysis of knee flexion angles (°) measured by each examiner on different days with the smartphone goniometric application

	M1 ± *SD*	*M2* ± *SD*	*t* paired *p*	*r*	R^2	ICC (95% CI)
E1	44.6 ± 11.1	42.1 ± 11.1	0.011	0.886*	0.785	0.860 (0.736–0.928)
E2	45.3 ± 10.1	43.3 ± 11.3	0.075	0.834*	0.721	0.828 (0.701–0.911)
E3	45.2 ± 10.6	44.2 ± 11.6	0.312	0.856*	0.733	0.851 (0.721–0.923)

E1 – first examiner, E2 – second examiner, E3 – third examiner, *M1* – mean of the first day, *M2* – mean of the second day, *SD* – standard deviation, *r* – Pearson's correlation coefficient, R^2 – coefficient of determination, ICC – intra-class correlation coefficient, 95% CI – 95% confidence interval
* $p < 0.010$



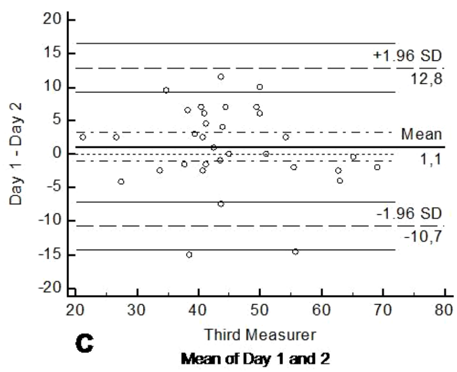


Figure 3. Bland-Altman dispersion plot for intra-examiner agreement for examiner 1 (A), examiner 2 (B), and examiner 3 (C)

Table 2. Inter-examiner analysis of knee flexion angles (°) measured by each examiner on different days with the smartphone goniometric application

		M ± *SD*	*M* ± *SD*	*t* paired *p*	*r*	R^2	ICC (95% CI)
Day 1	E1-E2	44.6 ± 11.1	45.3 ± 10.1	0.286	0.943*	0.889	0.966 (0.937–0.984)
	E1-E3	44.6 ± 11.1	45.2 ± 10.6	0.276	0.954*	0.910	0.954 (0.912–0.988)
	E2-E3	45.3 ± 10.1	45.2 ± 10.6	0.885	0.973*	0.947	0.986 (0.951–0.993)
Day 2	E1-E2	42.1 ± 11.1	43.3 ± 11.3	0.058	0.950*	0.903	0.975 (0.906–0.987)
	E1-E3	42.1 ± 11.1	44.2 ± 11.6	0.051	0.961*	0.924	0.980 (0.929–0.990)
	E2-E3	43.3 ± 11.3	44.2 ± 11.6	0.186	0.948*	0.899	0.973 (0.906–0.987)

E2 – second examiner, E3 – third examiner, *M* – respective mean of examiners, *SD* – standard deviation, *r* – Pearson's correlation coefficient, R^2 – coefficient of determination, ICC – intra-class correlation coefficient, 95% CI – 95% confidence interval
* $p < 0.010$

Table 3. Bias and upper and lower limits of agreement for inter-examiner results submitted to Bland-Altman analysis

		Bias	ULA	LLA	*p*
Day 1	E1-E2	–0.7	6.6	–8.0	0.290
	E1-E3	–0.6	5.9	–7.1	0.270
	E2-E3	0.0	5.0	–4.9	0.910
Day 2	E1-E2	–1.2	5.7	–8.1	0.060
	E1-E3	–2.0	4.3	–8.3	< 0.001*
	E2-E3	–0.9	6.4	–8.1	0.180

E2 – second examiner, E3 – third examiner, bias – mean difference between measures, ULA – upper limit of agreement (bias + 1.96*SD), LLA – lower limit of agreement (bias – 1.96*SD)
* $p < 0.050$ compared with 0

The Pearson's correlation coefficient indicated good intra-examiner correlation ($r > 0.80$; $p < 0.001$) and the mean coefficient of determination was 0.75, implying that 75% of knee flexion measured on the first day was explained by the variable analysed on the second day.

In the Bland-Altman analysis for intra-examiner results (Figure 3), the bias differed from 0 only for the first examiner ($p = 0.010$; Figure 3A), but the differences were less than 5° for all examiners. For the first examiner, 3% of the data exceeded the limits of agreement, whereas the proportion for examiners 2 (Figure 3B) and 3 (Figure 3C) was 5.8%. As the expected value was 5%, the results could be considered acceptable.

For the inter-examiner data (Table 2), the paired *t*-test revealed no significant differences among the examiners ($p > 0.050$). The inter-examiner reliability was considered excellent (ICC > 0.90) for all examiners ($p < 0.001$). The Pearson's correlation coefficient indicated high inter-examiner correlation ($r > 0.90$; $p < 0.001$) and the mean coefficient of determination was 0.91, implying that 91% of the degree of knee flexion measured by one examiner was explained by that measured by another examiner.

The Bland-Altman analyses indicated agreement among the data for five of the six analyses (bias: $p > 0.050$). Three of the six graphs demonstrated data exceeding 5% (data not shown), but the proportion was only 5.8% (Table 3).

Discussion

The aim of the present study was to evaluate the reliability of knee flexion measurements performed in healthy women with the use of the ROM© goniometric smartphone application. Three independent trained examiners carried out the evaluations and achieved satisfactory precision, as demonstrated by the results of both intra- and inter-examiner reliability analyses.

With regard to the intra-examiner reliability, the *t*-test revealed a difference of 2.5° between the first and second day for examiner 1. However, small angular differences do not have a clinically important impact on the evaluation of joint movement [9]. Some researchers suggest that the minimum clinically relevant measurable difference is from 5° [22] to 10° [24].

A strong point of the present study was the use of the agreement analysis proposed by Bland and Altman [21], which has recently been suggested as the best way to evaluate agreement among data [23], since correlation analyses quantify the degree by which two variables are related and may overestimate the results [25]. The Bland-Altman analysis demonstrated good intra-examiner agreement, with a bias lower than that defined *a priori* (5°), ranging from 1.1° for examiner 3 to 2.5° for examiner 1. Although some measurements exceeded the established limit, the difference was small (5.8° rather than 5°) and could therefore be considered acceptable [23]. The same was found for the inter-examiner measurements, as five of the six analyses demonstrated no difference ($p > 0.05$) when compared with 0, and the bias ranged from 0° to 2°, with the upper and lower limits of agreement spanning less than 8°.

The present results demonstrated excellent intra- and inter-examiner reliability, as well as good to high correlations among the measurements. Previous studies indicate a high degree of reliability regarding the angles of the knee when using smartphone goniometric applications in patients submitted to knee reconstruction surgery [3], the measurement of the angle of external tibial rotation [26], as well as comparisons between experienced and novice examiners [6] and among inexperienced examiners measuring maximum knee flexion in healthy individuals [9].

Reliability analyses have also been performed for other joints. Shin et al. [11], Alba-Martín [15], and Johnson et al. [16] found excellent intra- and inter-examiner reliability for measurements of the range of motion of the shoulder in different positions using the Clinometer© application. Ferriero et al. [10] revealed high intra- and inter-examiner reliability (ICC > 0.99) regarding angular measurements of the elbow with the DrGoniometer© application for the iPhone©. Charlton et al. [13] observed good to excellent reliability with reference to passive range of motion of the hip in 20 healthy men using the Hip ROM Tester© application. The excellent reliability demonstrated in the present study may be explained by the ease of use of the smartphone for measuring the range of motion of the knee between examiners as well as by the same examiner on different days. Likewise, Hambly et al. [9] found an ICC of 0.89 and the Pearson's correlation coefficient of 0.79 for the intra-examiner reliability using the iGoniometer application for the iPhone©.

A large number of applications are available for measuring joint range of motion. Some can be used for any joint, whereas others are dedicated for specific joints [27]. Applying similar methodology as that employed in the present study, Milanese et al. [6] evaluated a specific application for the knee: Knee Goniometer App©, KGA (Apple iPhone©). The present investigation is the first one to evaluate angular measurements of the knee with an application designed for any joint: ROM© for Android Samsung Galaxy S5.

With the aim of investigating differences among applications and operating systems, Kuegler et al. [27] compared the angular measurements of sixteen goniometric applications and two different operating systems using an irregular pyramid-shaped object with four sides and consequently four different angles. The applications performed in a similar manner for all measurements, with minimal variation; the iOS (operating system of Apple) had a maximum angular variation of 1.8°, and the Android (operating system of Google) presented a maximum angular variation of 1.3°. This demonstrates that the different operating systems used on Apple and Samsung and the different applications available exert little influence on the results of the angular measurements of an object. However, no studies have compared applications and operating systems for the evaluation of individuals.

The present study has limitations that should be addressed. The smartphone could have slipped out of position on the skin of the participant, since it was only held in place by the examiner. A support attached to the leg could reduce this problem, but would not adequately reflect clinical practice. The smartphone employed had less than one year of use and it is not possible to affirm that the angular measurements will remain reliable over a longer period of use because of the possibility of the device de-calibration. Finally, the findings can only be extrapolated to healthy young women and not to other populations.

On the basis of the procedures employed in the present study, the ROM© goniometric application for the smartphone can be used in clinical practice for evaluating the range of motion of the knee in healthy women, as demonstrated by the excellent intra- and inter-examiner reliability and good agreement.

References

1. Milani P, Coccetta CA, Rabini A, Sciarra T, Massazza G, Ferriero G. Mobile smartphone applications for body position measurement in rehabilitation: a review of goniometric tools. PM & R. 2014:6(11):1038–1043; doi:10.1016/j.pmrj.2014.05.003.
2. Tacconi C, Mellone S, Chiari L. Smartphone-based applications for investigating falls and mobility. 5th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2011, Dublin, Ireland, May 23–26, 2011; doi:10.4108/icst.pervasivehealth.2011.246060.
3. Jenny JY. Measurement of the knee flexion angle with a smartphone-application is precise and accurate. J Arthroplasty. 2013;28(5):784–787; doi:10.1016/j.arth.2012.11.013.
4. Duarte RB, Romanatti SV, Gongora H, Antonietti LS, Pires L, Cohen M. Reliability indices of popliteal angle analysis through photogrammetry [in Portuguese]. Rev Bras Med Esp. 2014;20(6):416–420; doi:10.1590/1517-86922014200601024.
5. César EP, Gomes PSC, Marques CL, Domingos BDP, Santos TM. Intra-rater reliability of knee flexion and extension range of motion measurement through the photogrammetry method [in Portuguese]. Fisioter Pesq. 2012; 19(1):32–38; doi: 10.1590/S1809-29502012000100007.
6. Milanese S, Gordon S, Buettner P, Flavell C, Ruston S, Coe D, et al. Reliability and concurrent validity of knee angle measurement: smartphone app versus universal goniometer used by experienced and novice clinicians. Man Ther. 2014;19(6):569–574; doi:10.1016/j.math.2014.05.009.
7. Santos CM, Ferreira G, Malacco PL, Sabino GS, Moraes GFS, Felício DC. Intra and inter examiner reliability and measurement error of goniometer and digital inclinometer use. Rev Bras Med Esp. 2012;18(1):38–41; doi:10.1590/S1517-86922012000100008.
8. Sacco ICN, Alibert S, Queiroz BWC, Pridas D, Kieling I, Kimura AA, et al. Reliability of photogrammetry in relation to goniometry for postural lower limb assessment. Rev Bras Fisioter. 2007;11(5):411–417; doi:10.1590/S1413-35552007000500013.
9. Hambly K, Sibley R, Ockendon M. Level of agreement between a novel smartphone application and a long arm goniometer for assessment of maximum active knee flexion by an inexperienced tester. Int J Physiother Rehabil. 2012;2:1–14.

10. Ferriero G, Sartorio F, Foti C, Primavera D, Brigatti E, Vercelli S. Reliability of a new application for smartphones (DrGoniometer) for elbow angle measurement. PM & R. 2011;3(12):1153–1154; doi:10.1016/j.pmrj.2011.05.014.
11. 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(4):298–304; doi: 10.1016/j.math.2012.02.010.
12. Werner BC, Holzgrefe RE, Griffin JW, Lyons ML, 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(11):275–282; doi:10.1016/j.jse.2014.02.030.
13. Charlton PC, Mentiplay BF, Pua YH, 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;18(3):262–267; doi:10.1016/j.jsams.2014.04.008.
14. Otter SJ, Agalliu B, Baer N, Hales G, Harvey K, James K, et al. The reliability of a smartphone goniometer application compared with a traditional goniometer for measuring first metatarsophalangeal joint dorsiflexion. J Foot Ankle Res. 2015;8(1):1–7; doi:10.1186/s13047-015-0088-3.
15. Alba-Martín R. Reliability and validity of shoulder and elbow measurements: analysis of an Android application and a goniometer [in Spanish]. Rehab. 2016;50(2):71–74; doi:10.1016/j.rh.2015.12.002.
16. Johnson LB, Sumner S, Duong T, Yan P, Bajcsy R, Abresch RT, et al. Validity and reliability of smartphone magnetometer-based goniometer evaluation of shoulder abduction – a pilot study. Man Ther. 2015;20(6):777–782; doi:10.1016/j.math.2015.03.004.
17. Santos RA, Derhon V, Brandalize M, Brandalize D, Rossi LP. Evaluation of knee range of motion: correlation between measurements using a universal goniometer and a smartphone goniometric application. J Bodyw Mov Ther. 2016, in press; doi:http://dx.doi.org/10.1016/j.jbmt.2016.11.008.
18. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions, 3rd ed. New York: John Wiley & Sons; 2003.
19. Rothstein JM, Miller PJ, Roettger RF. Goniometric reliability in a clinical setting. Elbow and knee measurements. J Pediatr Orthop. 1984;4(2):265.
20. Field A. Discovering statistics using SPSS, 4th ed. Thousand Oaks: Sage Publications; 2013.
21. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327(8476):307–310; doi:10.1016/S0140-6736(86)90837-8.
22. Chaudhary R, Beaupré LA, Johnston DW. Knee range of motion during the first two years after use of posterior cruciate-stabilizing or posterior cruciate-retaining total knee prostheses. A randomized clinical trial. J Bone Joint Surg Am. 2008;90(12):2579–2586; doi:10.2106/JBJS.G.00995.
23. Giavarina D. Understanding Bland Altman analysis. Biochem Med. 2015;25(2):141–151; doi:10.11613/BM.2015.015.
24. Jakobsen TL, Christensen M, Christensen SS, Olsen M, Bandholm T. Reliability of knee joint range of motion and circumference measurements after total knee arthroplasty: does tester experience matter? Physiother Res Int. 2010;15(3):126–134; doi:10.1002/pri.450.
25. Hirakata VN, Camey SA. Bland-Altman analysis of agreement between methods [in Portuguese]. Rev. HCPA. 2009;29(3):261–268.
26. Jeon I, Kwon O, Weon J, Ha S, Kim S. Reliability and validity of measurement using smartphone-based goniometer of tibial external rotation angle in standing knee flexion. Phys Ther Kor. 2013;20(2):60–68; doi:10.12674/ptk.2013.20.2.060.
27. Kuegler P, Wurzer P, Tuca A, Sendlhofer G, Lumenta DB, Giretzlehner M, et al. Goniometer-apps in hand surgery and their applicability in daily clinical practice. Saf Health. 2015:1–11; doi:10.1186/s40886-015-0003-4.