



Original Article

The reliability of goniometric measurement of frontal knee alignment in standing in people with knee osteoarthritis: A pilot study

Thiv Rodrigo, B Pty.¹, Andrew Hislop, PhD., B Pty.², Shaun P. O'Leary, PhD., B Pty.¹, Peter J. Window, PhD., B Pty.³¹School of Health and Rehabilitation Sciences, University of Queensland, St Lucia, ²Department of Physiotherapy, The Prince Charles Hospital, Chermside, ³Department of Physiotherapy, Royal Brisbane and Women's Hospital, Herston, Queensland, Australia.

*Corresponding author:

Peter James Window,
Department of Physiotherapy,
Royal Brisbane and Women's
Hospital, Herston, Queensland,
Australia.

peter.window@health.qld.gov.au

Received: 14 November 2024
Accepted: 23 January 2025
Epub ahead of print: 08 March 2025
Published: 02 April 2025

DOI

10.25259/JMSR_502_2024

Quick Response Code:



ABSTRACT

Objectives: Goniometry is a low-cost, high-access measure of frontal knee alignment in people with knee osteoarthritis (KOA). Previous literature has demonstrated the clinical relevance of measuring knee alignment; however, the reliability and associated measurement error of goniometer methods have not been adequately investigated. We sought to evaluate the reliability and measurement error of goniometric measurement of frontal alignment of the knee in standing individuals with KOA.**Methods:** An intra- and inter-rater reliability study design was conducted. Twenty-six individuals with KOA based on clinical and radiological diagnosis criteria were recruited at a tertiary hospital health service. Frontal knee alignment was measured by two examiners (Examiner A, an experienced physiotherapist, and Examiner B, a final-year pre-registration physiotherapy student). Each examiner independently recorded two measures. Reliability, measurement error, and limits of agreement were evaluated using Intraclass Correlation Coefficients (ICCs; 3, k), standard error of measurement (SEM), and Bland-Altman plots, respectively.**Results:** Excellent intra- (ICC = 0.99) and inter-rater reliability (ICC = 0.96) were demonstrated. Inter-rater SEM was 1.20°, while intra-rater SEM ranged between 0.58 and 0.59° for each examiner (MDC₉₅ Examiner A: 1.64°, MDC₉₅ Examiner B: 1.61°). Bland-Altman plots revealed LOA₉₅ of, at worst, 5.8° between examiners and 2.4° for individual raters.**Conclusion:** This study demonstrated excellent reliability (intra- and inter-rater) of goniometric frontal knee alignment measurement in people with KOA. However, to ensure high levels of reliability, measurement methods of alignment need to be standardized between practitioners so that clinical decisions can be confidently made based on this measure.**Keywords:** Alignment, Genu valgum, Genu varum, Goniometer, Knee, Osteoarthritis, Physical exam, Validation study

INTRODUCTION

Knee osteoarthritis (KOA) is a prevalent disease and one of the leading causes of disability globally.^[1] Current clinical practice guidelines for managing KOA highlight the importance of**How to cite this article:** Rodrigo T, Hislop A, O'Leary SP, Window PJ. The reliability of goniometric measurement of frontal knee alignment in standing in people with knee osteoarthritis: A pilot study. J Musculoskelet Surg Res. 2025;9:260-5. doi: 10.25259/JMSR_502_2024

land-based exercise and weight loss (when appropriate) as key elements of a non-surgical approach.^[2] It has a complex and multifactorial etiology, with genetic, anatomical, and mechanical determinants likely influencing disease onset and progression.^[3] Various patient characteristics may also impact non-surgical treatment response.^[4]

Frontal knee alignment has typically influenced surgical decision-making^[5] as well as exercise selection^[6] for individuals with KOA. In addition, frontal knee malalignment has been associated with a poor outcome following non-surgical care,^[4] increased likelihood of disease progression,^[7-9] and progression to total knee arthroplasty.^[10] Therefore, frontal knee plane alignment assessment appears clinically indicated to inform management decision-making in KOA.

Full limb radiographs are commonly accepted as the gold standard to assess frontal knee alignment, where the knee mechanical axis is calculated from the intersecting lines formed from the femoral axis (center of the femoral head to the center of the femoral intercondylar notch) and the tibial axis (center of the tibial spines to the center of the ankle).^[11,12] However, these are costly, require specialized equipment, and are not routinely collected as part of first-line management strategies. Goniometer measures of frontal knee alignment present a readily accessible and cost-effective method of measuring alignment to inform non-surgical decision-making. Studies exploring the relationship of goniometric measures to radiographic methods are mixed. A poor relationship ($r = 0.32$) was observed between a long-arm goniometer measure and the mechanical axis utilizing the patellar tendon below the knee as the distal landmark.^[13] Stronger associations have been demonstrated when the fixed arm of the goniometer is aligned to the center of the ankle, with correlations ranging from $r = 0.50$ to $r = 0.70$.^[11,14,15] However, the reliability of these measures has been poorly investigated. Previous literature has established reliability for knee alignment ranging from Intraclass Correlation Coefficient (ICC) 0.79–0.99. However, these studies have utilized landmarks that demonstrate a poor relationship to the mechanical axis^[13,16] and/or were limited to either intra- or inter-rater measures.^[11,13,14,16] In addition, no studies were identified that quantified the associated standardized error of the measurement (SEM), minimal detectable change (MDC), or limits of agreement (LOA).

This study aimed to establish the reliability, associated measurement error, and LOA of goniometric measurement of frontal knee alignment in standing in individuals with KOA. To establish the reliability across different experience levels of practitioners, we secondly aimed to compare measurement agreement between a final-year pre-registration physiotherapy student and an experienced physiotherapist. It was hypothesized that good excellent intra-rater and

inter-rater reliability (≥ 0.75) would be observed, with greater intra-rater reliability for the experienced physiotherapist.

MATERIALS AND METHODS

Study Design

An intra-and inter-rater reliability study was conducted, nested within a larger multi-site prospective longitudinal study investigating the validity of a clinical nomogram to predict the outcome of conservative multidisciplinary treatment of KOA.^[17] This study accorded with strengthening the reporting of observational studies in epidemiology (STROBE) and guidelines for reporting reliability and agreement studies (GRRAS) guidelines.^[18,19] Participants provided written informed consent.

Participants

Participants were recruited within a physiotherapy-led advanced practice orthopedic service in a tertiary hospital in Brisbane, Australia. Eligibility for participation was assessed by an advanced musculoskeletal physiotherapist and based on clinical diagnostic criteria for KOA,^[20] including persistent knee pain on most days for at least 1 month in the preceding year, morning stiffness limited to <30 min, functional limitation, and radiological signs of KOA (based on radiology report or clinician's judgment). Participants were excluded if they presented with potentially serious medical conditions (red flags), knee pain referred from the hip or back, active inflammatory conditions, were likely to be aggravated by conservative treatment due to severity of symptoms, significant neurovascular involvement, or if specialist medical consultation was specifically requested. In cases where both knees were affected, the most problematic/severe knee was assessed for the purposes of this study.

Sample size estimation

Based on previous goniometry reliability studies, a minimum sample of 20 participants was considered satisfactory to address the research aims.^[21,22] In addition, it was estimated that 18.4 participants would be sufficient to achieve significance ($P < 0.05$) with two raters (using 80% power) given minimally acceptable and ideal ICC values of 0.7 and 0.9, respectively.^[23] Additional participants were recruited to enhance the generalizability of findings while also accounting for potential incomplete data or participant dropout.

Goniometry measurement of frontal knee alignment

The measure of frontal knee alignment was performed in standing using a standard universal goniometer. The goniometer axis was placed over the center of the patellar, with the moving arm aligned to the anterior superior iliac spine



Figure 1: Goniometric measurement of frontal knee alignment with anatomical landmarks annotated. ASIS: Anterior superior iliac spine.

(ASIS) and the stationary arm aligned at the center of the ankle [Figure 1].^[4,11] Each examiner identified the respective landmarks through palpation. The measured knee angle was recorded as either neutral (0°), varus ($+$), or valgus ($-$).

Procedure

An experienced musculoskeletal physiotherapist with 16 years of clinical experience [Andrew Hislop (AH)] and a final-year pre-registration physiotherapy student with <1-year of clinical experience [Thiv Rodrigo (TR)] recorded the measurements. Before testing, both examiners practiced landmark palpation, goniometer arm placement, and angle measurement recording together to ensure familiarity with the method and standardization of the measurement procedure.

Frontal knee alignment measures were undertaken during the participant's routine initial consultation in the service. Each assessor recorded two measurements and was blinded to each other's measurements by leaving the examination room. Participants were asked to march on the spot before putting equal weight through their legs before the first measurement was recorded. Participants were instructed to refrain from using external mobility aids during this process. Measurements were obtained in the same session to reflect clinical applicability, and examiners were not blinded to their initial measurement. However, they were instructed not to look at the goniometer measurement until they were satisfied that the landmarks were accurately aligned. The second measurement by each examiner was taken 5 minutes following the first measurement. The initial assessor then exited the room, and the second assessor repeated the process. To reduce measurement bias, the order of testing between assessors was randomized.

Statistical analysis

Descriptive statistics (group mean, standard deviation, and 95% confidence intervals) were computed. ICC and SEM (absolute reliability) indices were used to calculate the reliability coefficients (intra- and inter-rater) for the repeated knee angle measurement.^[24] As the same raters were used to determine reliability, ICC values were determined from a two-way mixed effects model (ICC 3, k)^[25,26] with absolute agreement and interpreted as poor (<0.5), moderate ($0.5-0.75$), good ($0.75-0.9$), or excellent (>0.90) reliability.^[27] Each examiner's two measures were used to calculate their intra-rater reliability, while inter-rater reliability was determined using the mean of each examiner's measurements. The SEM was calculated using the formula: $SEM = \text{Standard Deviation} \times (\sqrt{1 - ICC})$.^[28] The MDC_{95} was calculated using the formula: $MDC = SEM \times 1.96 \times \sqrt{2}$.^[29] Paired samples t -tests were used to evaluate examiner scores for systematic differences. Bland-Altman plots with 95% LOA_{95} were used to provide an indication of variation between testers, with the experienced examiner (Examiner A) deemed the reference standard. Analyses were conducted using the IBM Statistical Package for the Social Sciences Statistics 26 (IBM, New York, USA).

RESULTS

Twenty-six adults (10 females) (mean [standard deviation (SD)] age of 62.3 [9.8] years) participated, with a mean (SD) body-mass index (BMI) of 32.4 (6.79) kg/m^2 . Participants, on average, had a varus alignment (experienced physiotherapist 2° , physiotherapy student 0.9°), with a measurement range of -13° (valgus) to 16° (varus).

Reliability between examiner A and B

Reliability findings (inter- and intra-rater) are summarized in Table 1. Inter-rater reliability was excellent (ICC 3, k : 0.96) with an SEM of 1.20° . Intra-rater reliability of both examiners was near identical (ICC 3, k : 0.99), with calculated SEM ranging from 0.58° to 0.59° .

The paired samples t -test demonstrated a small but significant systematic difference between raters (mean difference 1.13° , $P = 0.02$), with examiner A recording scores that were slightly higher in the varus direction. The Bland-Altman plot of agreement between both examiners is seen in Figure 2, 95% LOA ranged from -3.52° to 5.79° .

DISCUSSION

This study sought to establish the within-session reliability and associated measurement error of a frontal knee alignment measure using a goniometer in people with KOA. Overall, the goniometer method demonstrated excellent levels of

Table 1: Inter-rater and intra-rater reliability.

Variables	ICC (3,k)	95% CI of ICC (3, k)	P-value	SEM (°)	MDC (°)
Inter-rater reliability	0.96	0.91–0.98	<0.05	1.20	3.33
Intra-rater reliability					
Examiner A	0.99	0.98–0.99	<0.05	0.59	1.64
Examiner B	0.99	0.98–0.99	<0.05	0.58	1.61

The ICC (3, k) and accompanying 95% confidence interval ranges are provided with the SEM and MDC values. P-value denotes the significance of the ICC (3, k), with values <0.05 representing significance. Examiner A (experienced Physiotherapist). Examiner B (Physiotherapy student). ICC: Intraclass correlation coefficients, SEM: Standard error of the measurement, MDC: Minimal detectable change, CI: Confidence interval

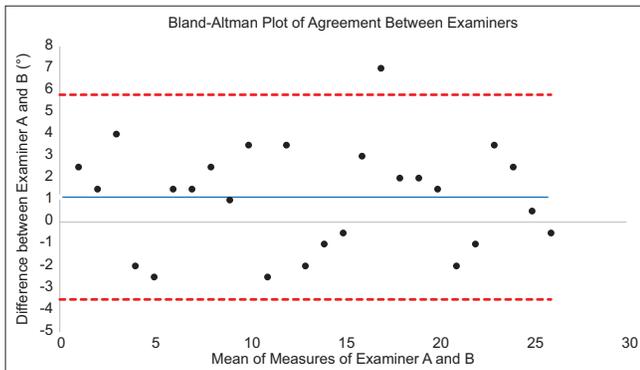


Figure 2: Bland–Altman scatterplot for agreement between Examiner A and B. Black dots represent the difference in measurements between examiners for individual participants, while the blue line represents the mean difference. Upper and lower limits of agreement (ULO and LLO) are represented by the upper and lower dashed lines, respectively.

reliability, including inter-rater (between an experienced physiotherapist and physiotherapy student) and intra-rater reliability. This, combined with small SEM values and LOA that were at worst 5.79°, suggests that the goniometry method described in this study is suitable for clinical use both between and within assessors.^[30]

The results obtained in this study are consistent with other studies investigating the reliability of frontal plane knee alignment in similar populations.^[13,14,16] The reliability of frontal knee alignment appears to be higher when the measurement procedure is standardized,^[16] which was reflected in our study's excellent agreement. However, it is noted that a small but statistically significant systematic bias (1.13°) was observed, with the more experienced examiner generally recording values that were slightly more varus in direction. We also utilized averaged measures for assessing inter-rater reliability that may lead to higher agreement and should be considered by clinicians performing the measurement. Excellent intra-rater reliability was seen irrespective of experience level, consistent with other studies demonstrating no significant differences in goniometry reliability across practitioners of varying skill levels.^[21,31]

This study aligned the goniometer axis to the center of the patella, aligning the two arms to the ASIS superiorly and the

center of the ankle inferiorly, consistent with other studies.^[4,11] As the arms of the goniometer do not reach the superior and inferior landmarks in this method, visual alignment by each examiner is required and some measurement error was anticipated. Further, measurement error was expected given our cohort's high average BMI (32.4kg/m²), with increased weight a factor increasing discrepancy when palpating the ASIS.^[32] Despite this, calculated reliability was greater than that of a previously described method utilizing the patellar tendon and mid-thigh as landmarks.^[13] This may indicate that the potential error introduced through visual alignment and/or increased adiposity is offset using landmarks at a greater distance from the axis center. The error introduced through palpation or malalignment translates to smaller angular changes as the distance from the axis center increases.

From a clinical perspective, the findings of this study indicate acceptable consistency with a between-rater SEM of 1.20°. O'Leary *et al.* found a greater degree of knee varus to be associated with a poorer outcome following conservative treatment of KOA. Specifically, those with a poor response had an average frontal knee varus angle of 1.84°, while those with a good response had a valgus angle of 2.11°.^[4] Although the SEM coefficient observed in this present study is smaller than this responder difference (total 3.95°), the upper LOA between examiners (5.79°) highlighted potential challenges in discriminating response in individuals with small differences in frontal alignment, which would introduce some uncertainty and would require consideration in clinical decision-making. However, these limits are broadly consistent with 5° cutoff values predictive of deteriorating physical function and pain in KOA when evaluating knee malalignment using full limb radiographs.^[12] This potential uncertainty would also be reduced in the case of a single examiner, with SEMs of 0.58–0.59°. A calculated MDC of 3.33° (1.61–1.64° for within raters) provides additional information regarding the utility of the goniometer method for undertaking repeated measures over time to evaluate potential disease progression.

This study has strengths and limitations. It was undertaken in a clinical environment within the constraints of normal clinical time, in a cohort with a high BMI, enhancing the

generalizability of the findings to a clinical context. However, examiners were not blinded to their initial measurement, and these were completed within-session, leading to potential overestimation of the intra-rater reliability due to bias from examiners recording their second measurement close to their first. However, the risk was mitigated by requesting the examiners to avoid looking at the goniometer measurement until the final alignment had been determined. While adequately powered, the study also had a relatively low sample size and future studies should seek to confirm the findings on a larger scale, including subgroup analyses such as sex and BMI, within clinical populations with other lower limb conditions, and including between-session evaluations of reliability.

CONCLUSION

This study established that frontal knee alignment measured using a universal goniometer in individuals with KOA has excellent intra- and inter-rater reliability. The findings suggest that the method described in this study has potential clinical application when used by different practitioners. Some measurement errors were demonstrated and should be considered when making clinical decisions based on this measure.

Recommendation: Goniometer-based assessment of frontal plane knee alignment in standing provides reliable measurements in clinical settings with small measurement errors. Some care should be taken when interpreting small differences between raters. Further research should be conducted to evaluate these findings in a larger sample.

Acknowledgment: The authors thank the participants for their involvement.

Authors' contributions: TR: Contributed to the study design, completed data collection, undertook data analysis, and wrote the initial draft of the article. AH: Contributed to the study design, completed data collection, and reviewed the draft manuscript. SO: Contributed to the study design and assisted with drafting and revising the manuscript. PW: Contributed to the study design, assisted with data analysis, and assisted with drafting and revising the initial manuscript. All authors have critically reviewed and approved the final draft and are responsible for the manuscript's content and similarity index.

Ethical approval: This study was conducted in accordance with the Declaration of Helsinki and received ethics approval through the Royal Brisbane and Women's Hospital Ethics committee May on 28th, 2021. Ethics Approval Number: HREC/2019/QRBW/58669.

Declaration of participant consent: The authors certify that they have obtained all appropriate participant consent forms. In the form, the participants have given their consent for their images and other clinical information to be reported in the journal. The participants understand that their names and initials will not be published, and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.

Use of artificial intelligence (AI)-assisted technology for manuscript preparation: The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using the AI.

Conflicts of interest: There are no conflicting relationships or activities.

Financial support and sponsorship: This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Hawker GA. Osteoarthritis is a serious disease. *Clin Exp Rheumatol* 2019;37 (Suppl 120):3-6.
- Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. East Melbourne, Victoria: RACGP; 2018.
- Weiss E, Jurmain R. Osteoarthritis revisited: A contemporary review of aetiology. *Int J Osteoarchaeol* 2007;17:437-50.
- O'leary S, Raymer M, Window P, Swete Kelly P, Elwell B, Mcloughlin I, *et al.* Patient characteristics associated with a poor response to non-surgical multidisciplinary management of knee osteoarthritis: A multisite prospective longitudinal study in an advanced practice physiotherapist-led tertiary service. *BMJ Open* 2020;10:e037070.
- Macedessi SJ, Allom RJ, Griffiths-Jones W, Chen DB, Wood JA, Bellemans J. The importance of joint line obliquity: A radiological analysis of restricted boundaries in normal knee phenotypes to inform surgical decision making in kinematically aligned total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2022;30:2931-40.
- Lim BW, Hinman RS, Wrigley TV, Sharma L, Bennell KL. Does knee malalignment mediate the effects of quadriceps strengthening on knee adduction moment, pain, and function in medial knee osteoarthritis? A randomized controlled trial. *Arthritis Care Res* 2008;59:943-51.
- Amaratunga HA, Adikari SB, Dassanayake TL, Gamage J, Suraweera HJ. Relationship between the goniometric alignment and articular cartilage damage in knee osteoarthritis. *Ceylon Med J* 2017;62:167-74.
- Brouwer GM, Tol AW, Bergink AP, Belo JN, Bernsen RM, Reijman M, *et al.* Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. *Arthritis Rheum* 2007;56:1204-11.
- Tanamas S, Hanna FS, Cicuttini FM, Wluka AE, Berry P, Urquhart DM. Does knee malalignment increase the risk of development and progression of knee osteoarthritis? A systematic review. *Arthritis Rheum* 2009;61:459-67.
- Turcotte JJ, Kelly ME, Aja JM, King PJ, Macdonald JH. Risk factors for progression to total knee arthroplasty within two years of presentation for knee osteoarthritis. *J Clin Orthop Trauma* 2021;16:257-63.
- Navali AM, Bahari LS, Nazari B. A comparative assessment of alternatives to the full-leg radiograph for determining knee joint alignment. *Sports Med Arthrosc Rehabil Ther Technol* 2012;4:40.

12. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA* 2001;286:188-95.
13. Hinman RS, May RL, Crossley KM. Is there an alternative to the full-leg radiograph for determining knee joint alignment in osteoarthritis? *Arthritis Care Res* 2006;55:306-13.
14. Kraus VB, Vail TP, Worrell T, Mcdaniel G. A comparative assessment of alignment angle of the knee by radiographic and physical examination methods. *Arthritis Rheum* 2005;52:1730-5.
15. Riddle DL. Validity of clinical measures of frontal plane knee alignment: Data from the Osteoarthritis Initiative. *Manual Ther* 2012;17:459-65.
16. Cibere J, Bellamy N, Thorne A, Esdaile JM, Mcgorm KJ, Chalmers A, *et al.* Reliability of the knee examination in osteoarthritis: effect of standardization. *Arthritis Rheum* 2004;50:458-68.
17. Window P, Raymer M, Mcphail SM, Vicenzino B, Hislop A, Vallini A, *et al.* Prospective validity of a clinical prediction rule for response to non-surgical multidisciplinary management of knee osteoarthritis in tertiary care: A multisite prospective longitudinal study. *BMJ Open* 2024;14:e078531.
18. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *BMJ* 2007;335:806-8.
19. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, *et al.* Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud* 2011;48:661-71.
20. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, *et al.* EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;69:483-9.
21. Milanese S, Gordon S, Buettner P, Flavell C, Ruston S, Coe D, *et al.* Reliability and concurrent validity of knee angle measurement: Smart phone app versus universal goniometer used by experienced and novice clinicians. *Manual Ther* 2014;19:569-74.
22. Schlager A, Ahlqvist K, Rasmussen-Barr E, Bjelland EK, Pingel R, Olsson C, *et al.* Inter- and intra-rater reliability for measurement of range of motion in joints included in three hypermobility assessment methods. *BMC Musculoskelet Disord* 2018;19:376.
23. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Statist Med* 1998;17:101-10.
24. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005;19:231-40.
25. Mcgraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30-46.
26. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420-8.
27. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-63.
28. Denegar CR, Ball DW. Assessing reliability and precision of measurement: An introduction to intraclass correlation and standard error of measurement. *J Sport Rehabil* 1993;2:35-42.
29. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-60.
30. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: A practical guide. Cambridge, United Kingdom: Cambridge University Press; 2011.
31. Pereira LC, Rwakabayiza S, Lécurveux E, Jolles BM. Reliability of the knee smartphone-application goniometer in the acute orthopedic setting. *J Knee Surg* 2017;30:223-30.
32. Moriguchi CS, Carnaz L, Silva LC, Salasar LE, Carregaro RL, Sato TO, *et al.* Reliability of intra- and inter-rater palpation discrepancy and estimation of its effects on joint angle measurements. *Man Ther* 2009;14:299-305.