European Journal of Physiotherapy

Title Page

Manuscript Title: Reliability of knee joint position sense measurement: a comparison between goniometry and image capture methods.

Authors: Fiona Irving[1], Joseph Russell[2], Toby Smith[3]

Affiliations:

1. Physiotherapy Department, United Lincolnshire Hospitals Trust, Pilgrim Hospital, Boston, United Kingdom.

2. Physiotherapy Department, Allied Health Professionals Suffolk, Hartismere Hospital, Eye, Suffolk, United Kingdom.

3. School of Health Sciences, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, United Kingdom.

Corresponding Author: Dr Toby Smith, School of Health Sciences, Queen's Building, Faculty of Medicine and Health Sciences, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom. Telephone: 044 (0)1603 593087; Fax: 044 (0)1603 593166; Email: toby.smith@uea.ac.uk.

Abstract

Aims: Evaluate the intra-rater and inter-rater reliability of hand-held goniometry compared to image capture (IMC) in the assessment of joint position sense (JPS) in healthy participants.

Methodology: Repeated-measures observational study design was undertaken with 36 asymptomatic university students of both genders aged between 18 to 45 years. JPS in the knee was assessed by two assessors over two sessions (one-week interval) using hand-held goniometry and IMC methods. Joint position sense was assessed at four target knee flexion angles. Intra- and inter-rater reliability was assessed with absolute error (AE), relative error (RE) and intra-class correlation coefficient.

Findings: Inter-rater reliability for goniometry was poor to substantial (ICC: 0.00 to 0.64) and was poor to moderate (ICC: 0.00 to 0.47) for IMC. Intra-rater reliability for goniometry was poor to moderate (ICC: 0.00 to 0.42) and poor to moderate for IMC (ICC: 0.00 to 0.41). AE for goniometry ranged from 3.2° to 8.6°, with RE from 0.1°-8.3°. For IMC, AE for goniometry was 5.3° to 12.5°, with RE ranging from 0.1° to 11.1°.

Principal Conclusions: Neither goniometry nor IMC appeared superior to the other in JPS assessment. Caution should be made when considering the reliability for goniometry and IMC before clinical assessment is made.

Keywords: Measurement; knee; proprioception; angle; range of motion

Introduction

Proprioception is the awareness of movement and position of a joint in space [1]. Proprioception relies on sensorimotor receptors which provide sensory input through visual, tactile and vestibular feedback systems [2,3]. Proprioception is also informed through motion where mechano-sensitive proprioceptors generate a feedback sensation which enables perceptual awareness of the limb including its movement, orientation in space, velocity, force and joint position sense (JPS) [1, 4-6].

Two previously documented methods of assessing JPS are hand-held goniometry and photographic image capture (IMC) [4]. Joint position sense is assessed through goniometry by positioning the joint under investigation at a pre-specified 'target'; angle as measured using the goniometer, asking the individual to try to remember that position, and then moving them out of that position, to then asking them to replicate the target angle, and re-measuring this angle. The same principle holds for IMC, where the target angle is measured with a goniometer and a photograph is taken of that joint angle. The participant then tries to remember that angle, moved out of that position and then replicates the target angle where a photograph is taken. The assessor then measures that angle of the repeated joint position to estimate the degree of agreement of deviation from the target angle.

Currently, limited evidence exists in relation to intra- and inter-rater reliability of IMC methods for JPS assessment [4]. Smith et al's [4] systematic review of JPS measures of the knee, suggested variable inter-rater and intra-rater reliability for IMC, but was unable to identify any studies which have assessed the reliability of hand-held goniometry in relation to knee JPS assessment. Ascertaining this is arguably a high priority given the importance of proprioception of the knee for everyday functional activity and its common association with injury and pathology [7-9]. Furthermore, given its proven reliability in knee range of movement (ROM) assessment [10] and its frequent use in clinical practice, hand-held goniometry clearly warrants further investigation which provided the rationale for this research study. Similarly, given the low cost and simplicity of JPS via digital photography measured by a protractor, this could be deemed the most appropriate and feasible method of IMC available for clinical practice [11-13]. This assertion, coupled with the lack of current research underpinning reliability of IMC techniques, has provided further rationale for digital photography IMC (referred to simply as IMC from hereon in) use in this study.

Accounting for this paucity of evidence on the reliability of goniometry JPS assessment, and since there is no previous evidence comparing goniometry to IMC, the purpose of this study was to evaluate the intra-rater and inter-reliability of hand-held goniometry compared to IMC assessment of JPS.

Methods

A repeated-measures design with two assessors was used to assess both intra- and inter-rater reliability.

Participants

Comparison of inter- and intra-observer reliability of goniometry and IMC in JPS measurement had not been previously assessed to base a sample size calculation on. It has been proposed that a minimum of 15 to 20 participants is necessary when determining the reliability of a quantitative variable [14]. Accordingly, accounting for the research timetable, 36 participants were recruited in total.

Participants were university students enrolled on either Physiotherapy or Occupational Therapy courses. Participants were recruited between November 2013 and January 2014. Twenty-seven participants were female and nine male aged 18 to 45 years (mean ± standard deviation age; 25.4 ± 6.0 years).

Participants were excluded if: they reported self-reported joint pain (any part of the body) experienced over the past three months; individuals allergic to adhesive tape; individuals who did not provide informed written consent; or were unable to undertake the entire assessment process.

Instrument and test procedure

Prior to testing, both assessors (Assessor 1; Assessor 2) were taught a standardised method of assessing JPS through goniometry and JPS methods as stated below. Both assessors had 12 months academic/clinical experience and were enrolled on a United Kingdom pre-registration physiotherapy masters-degree programme. This was taught by the chief investigator to ensure accurate and consistent with current specifications [15, 16]. Data collection was only commenced only once each assessor and the chief investigator were satisfied with the techniques adopted in accordance with the standardised techniques.

Joint position sense assessment was performed on the right knee of all participants to ensure consistency and prevent any potential variability in left and right JPS confounding the finding [17, 18]. Participants were prepared for assessment with application of white adhesive markers on the right greater trochanter, lateral tibiofemoral joint line and lateral malleolus (**Figure 1**). All testing was performed in standing with a 12cm distance between medial malleoli. Assessment of JPS was conducted in the following stages:

- Participant instructed by assessor to actively flex knee to first specified angle, termed the “target” angle. This angle was measured as per Norkin and White’s (1995) recommended methods to assess knee flexion, with a 15cm two-armed plastic hand-held goniometer with 1° marked increments [16].

- Participant instructed to remain in this angle of flexion for 10 seconds, to remember their knee position.

- Participant then instructed by assessor to straighten knee. Immediately following this, the participant was instructed to replicate the ’target’ angle position.

- In the goniometry method, this angle produced by the participant was measured using the goniometer whereas in the IMC assessment digital images were captured with a standard iPad 2 (model; A1395). The distance of the iPad to the limb ranged from 80 to 100 cm dependent on the length of the participant’s lower limb. These images were printed and the knee positions were measured using a simple 180° protractor.

- This process was repeated to assess four knee flexion target angles (20°, 40°, 75°, 100°).

Each assessor assessed each participant across the four target angles. Each measurement angle was recorded a single time. The order in which assessors evaluated participants was randomised through a single toss of a coin. A second coin toss determined the order of the two JPS measures (goniometry or IMC) to minimise the risk for order effects. All participants and assessors were instructed to be quiet throughout the assessment periods. Testing was performed in the same building throughout, in similar practical/clinical rooms, to reduce possible environmental variability.

All participants returned after a one-week interval, at the same time of day, when the process was repeated.

Statistical analysis

Descriptive statistics of mean and standard deviation assessed gender and age. JPS accuracy was measured by calculating absolute error (AE) and relative error (RE) [19]. AE was measured as the actual numerical difference between the test (target angle) and response angle recorded by the assessor [19]. Relative error (RE) was defined as the numerical difference between the test (target angle) and response angle (knee range of motion actually achived by the participant) with consideration of positive (overestimation) and negative (underestimation) values, represented as +/ - figures, therefore considering directional bias [19]. Both AE and RE were therefore necessary to determine the overall measurement error [20,21].

It was necessary to determine agreement between all AE and RE variables. Therefore the intra-class correlation coefficient (ICC) with 95% confidence intervals was selected to ascertain intra- and inter-rater reliability [17, 22]. Level of agreement strength for ICC's were categorised in boundaries as outlined by Landis and Koch (1977) [23]. Through this, values of less than 0.00 equate to poor strength, 0.00 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial and 0.81 to 1.00 as almost perfect.

All statistical analyses were completed using SPSS (version 21.0) (IMB, New York, USA).

Results

Intra-rater reliability of goniometry compared to image capture methods

Agreement strength for both AE and RE remained within two specified ICC boundaries for goniometry and IMC which did not exceed ‘moderate’. These were predominately distributed between ‘poor’ to ‘fair’ (Table 1; Table 2). Minimal differences were observed between AE and RE between variables in each grouping. Agreement for RE ranged between ‘slight’ to ‘moderate’ with ICC values achieving 0.00 (95% CI: 0.00 to 0.26) to 0.56 (95% CI: 0.29 to 0.74) (Table 1; Table 2). The largest AE and RE observed using goniometry assessment were 8.64° and 8.31º respectively which occurred at 40° for Assessor 2. The smallest AE and RE were 3.19° and -0.11° at 75° for Assessor 1 (Table 2).

Inter-rater reliability of goniometry compared to image capture method

Overall, the agreement of both assessors in goniometry and IMC categories of sessions one and two varied considerably across AE and RE. Range of ICC were between -0.00 (95% CI: 0.00 to 0.23) to 0.64 (95% CI: 0.00 to 0.31). This equates to ‘poor’ to ‘substantial’ agreement (Table 3). Results between the sessions remained within two ICC boundaries with the exception of AE values for session 1, which ranged from ‘poor’ to ‘substantial’ with ICC ranging 0.00 (95% CI: 0.00 to 0.20) to 0.64 (95% CI: 0.00 to 0.31). The greatest agreement of session one for both JPS methods was observed at 40° for Assessor 1 and 2 for goniometry. The lowest agreement was also observed at the same angle for IMC between the assessors for session one (Table 3). The most consistent agreement observed for both goniometry and IMC occurred at 100° for session two. In this instance, RE ICC values were both 0.47 (95% CI: 0.18 to 0.69) (Table 3). During IMC assessment, the largest AE and RE occurred for Assessor 2 at 100° during session two at values of 12.53° and -11.08° respectively. The smallest error occurred for Assessor 1 at 40° during session one at ICC of 5.36º (AE) and 0.08º (RE) (Table 3).

Overall findings of AE and RE of both methods for both assessors demonstrate greater error for IMC. Overall average error (standard deviation) of AE for goniometry was 5.16 (4.30) and 1.61 (6.50) for RE. The IMC resulted in AE of 8.20 (6.33) and -1.92 (10.18) for RE.

Intra-rater reliability of goniometry

Agreement between Assessor 1 and 2 ranged from ‘poor’ to ‘moderate’ for both AE and RE ICC (Supplementary Table 1). Agreement strength for both AE and RE values across all groupings were within two ICC boundaries. There was little difference observed between AE and RE for all groupings. Although definitively strong agreement was not observed overall, Assessor 2 demonstrated greater agreement in comparison to Assessor 1 with AE ICC between 0.05 (95% CI: 0.00 to 0.36) to 0.42 (95% CI: 0.12 to 0.67) (slight-moderate). In this instance RE ICC values achieving 0.24 (95% CI: 0.00 to 0.52) to 0.29 (95% CI: 0.02 to 0.56) (fair). Assessor 1 resulted in AE at 0.00 (95% CI: 0.00 to 0.23) to 0.26 (95% CI: 0.00 to 0.54) (poor-fair) and RE ICC between 0.00 (95% CI: 0.00 to 0.21) to 0.28 (95% CI: 0.00 to 0.55) (poor-fair) (Supplementary Table 1).

Inter-rater reliability of goniometry

Overall, agreement within this category highlighted wider inconsistency ranging from 0.00 (0.00 to 0.21) (poor) to 0.64 (95% CI: 0.00 to 0.31) (substantial) across AE and RE ICC (Supplementary Table 2). Agreement between assessors in session one showed greater strength overall in comparison to session two. In this case AE ICC were 0.00 (95% CI: 0.00 to 0.21) to 0.64 (95% CI: 0.00 to 0.31) (poor-substantial) and RE ICC between 0.00 (95% CI: 0.00 to 0.33) to 0.24 (95% CI: 0.00 to 0.50) (poor-fair) (Supplementary Table 2). Results at session two illustrated ICC for AE at 0.00 (95% CI: 0.00 to 0.23) to 0.34 (95% CI: 0.04 to 0.59) (poor-fair). The ICC values for RE were 0.09 (95% CI: 0.00 to 0.35) to 0.46 (95% CI: 0.17 to 0.69) (slight-moderate). However, as outlined above, session one demonstrated larger differences between AE and RE ICC resulting in greater difference in agreement (Supplementary Table 2).

The greatest agreement for RE in this group was during session two at 75°. Assessor 1 was-0.28º and Assessor 2 was -1.50º resulting in an ICC of 0.46 (moderate) (Supplementary Table 2). The weakest agreement for AE observed during session two at 40° with values of 3.33° for Assessor 1 and 7.31° for Assessor 2. The ICC value in this instance was 0.00 (95% CI: 0.00 to 0.23). The weakest agreement for

RE was observed in session one at 75° with values of -0.11º for Assessor 1 and -1.28º for Assessor 2. The ICC was 0.00 (95% CI: 0.00 to 0.33) (Supplementary Table 2).

Intra-rater reliability of IMC

Overall, the agreement observed was within two ICC boundaries. This did not exceed higher than a 'moderate' interpretation (Supplementary Table 3). The results demonstrate strong agreement was not observed across both sessions for either assessor. However AE of Assessor 1 demonstrate slightly higher agreement overall for all groupings comapred to Assessor 2. The ICC were 0.00 (95% CI: 0.00 to 0.26) to 0.41 (95% CI: 0.09 to 0.66) (poor-moderate) against ICC between 0.00 (95% CI: 0.00 to 0.10) to 0.32 (95% CI: 0.00 to 0.58) (poor to fair) (Supplementary Table 3). AE ICC for Assessor 2 were 0.00 (95% CI: 0.00 to 0.10) to 0.32 (95% CI: 0.00 to 0.58) (poor-fair). RE were 0.11 (95% CI: 0.00 to 0.41) to 0.43 (95% CI: 0.13 to 0.66) (slight-moderate). This highlighed considerable difference between AE and RE (Supplementary Table 3).

Inter-rater reliability of IMC

Agreement was within two ICC boundaries consistently across both sessions with minimal difference observed between AE and RE (Supplementary Table 4). Although high agreement was not observed between assessors across either session, stronger agreement was noted in session two. In this instance, ICC values for AE ranged from 0.03 (95% CI: 0.00 to 0.35) to 0.47 (95% CI: 0.18 to 0.69) (slight-moderate). ICC value for RE ranged from 0.15 (95% CI: 0.00 to 0.44) to 0.47 (95% CI: 0.18 to 0.69) (slight-moderate). ICC for AE and RE in session one were 0.00 (95% CI: 0.00 to 0.20) to 0.33 (95% CI: 0.01 to 0.59) (poor-fair) and 0.02 (95% CI: 0.21 to 0.28) to 0.30 (95% CI: 0.00 to 0.57) (slight-fair) respectively (Supplementary Table 4).

The strongest agreement of 'moderate' was observed during session two for 100°. This resulted in AE of 10.6° for Assessor 1 and 12.5° for Assessor 2 and an ICC of 0.47. RE values for this angle at session two are also consistent in 'moderate' agreement with -9.5° for Assessor 1 and -11.1° for Assessor 2, with an ICC of 0.47. The weakest agreement (poor) was observed during session one at 40° which resulted in AE of 5.4° for Assessor 1 and 7.6° for Assessor 2 with an ICC of 0.00 (95% CI: 0.00 to 0.20). RE for 40°

during session one resulted in 0.01° for Assessor 1 and 6.4° for Assessor 2 with an ICC of 0.02 (95% CI: 0.21 to 0.28). Accordingly, the strength of agreement for RE was 'slight' (Supplementary Table 4).

Discussion

The aim of this was to evaluate the intra-rater and inter-rater reliability of hand-held goniometry compared to image capture (IMC) in the assessment of joint position sense (JPS) in healthy participants. Clinically, establishing proprioceptive acuity is of high importance, given that proprioception plays a significant role in everyday functioning, joint stability, injury prophylaxis, and prevention of falls [3,6,21,24]. This demonstrates the necessity for establishing techniques that enable accurate measurement of proprioception through JPS for clinicians to identify individuals at risk of sustaining injury through proprioceptive deficit, objectively monitor pathological decline and to enable creation of specific rehabilitation programmes that both maintain and enhance proprioception in pathological and non-pathological populations [3,10, 21]. Therefore, overall further evidence is clearly warranted to determine the most reliable and accurate method of JPS assessment; following recent research development, emerging techniques such as smartphone applications could offer innovative and easily applicable approaches for clinical practice [25,26].

The largest AE and RE and consequently the greatest underestimation of a target angle occurred at 100° in this study which is the position most likely to cause fatigue for participants as research has suggested the chemical composition of muscle changes through fatigue, leading to irregularity of sensory output and increased joint laxity [27]. Several authors have proposed this secondary increase in laxity combined with temporary inefficiency of muscle receptors through fatigue, contributes to reductions in proprioceptive acuity and JPS accuracy which could account for the observation of these findings at this specific joint angle [27, 28]. Previous evidence has used three to five second holds whereas this study used ten seconds to sufficiently attempt JPS, which could have attributed to fatigue and consequently underestimation [13, 19, 29].

Joint position sense accuracy was not seen to improve towards end range movements despite some theories that would predict this to be the case due to increased articular compression, recruitment of mechanoreceptors thus leading to greater proprioceptive feedback and enhancing accuracy [20, 30]. However, this finding may be observed at extreme range of motion, but as producing extreme knee flexion may pose difficulty and most rehabilitation protocols utilise closed chain activity in the functional range of 0 to 100° in practice, it is arguably not appropriate to test such angles [31]. These findings could indicate that more accurate JPS assessment can be established at mid-joint range as supported by Barrack et al. in symptomatic populations [32].

Overall, the greatest overestimation of a target angle occurred at 20°. In standing, producing a 20° knee bend requires minimal flexion and thus potentially this small angle could be easier for participants to overestimate [20]. However, during bilateral weight-bearing, it has been reported that increased afferent input from all weight-bearing joints and other sensorimotor mechanisms influence and facilitate proprioceptive feedback and thus, AE and RE findings for knee JPS may be due to factors external to proprioceptive acuity at the knee joint [3,19].

It is critical when interpreting the results from this study that as participant variability and assessor measurement error cannot be separately examined in this study through the use of JPS as a measure of proprioception, it cannot be definitively ascertained if measurement error resulted from either one or a combination of the two. Although effects of fluctuations in an individual's circadian rhythm were controlled for where possible through completion of testing at similar timings for both sessions [33], uncontrollable factors such as behaviour of participant, individual physiology and learned effects could have affected accuracy and overall results which is a fundamental limitation of the study and limits the generalisability of the findings. It is also critical to consider that the average AE and RE values of all participants were reported in this study and while the mean (average) is widely used measure of central tendency, it can be susceptible to influence by outliers and thus caution should be adopted when interpreting the findings [17].

Goniometry has been routinely employed in clinical practice for many years, while the findings from this study in isolation are not sufficient to recommend deterring its use, they do highlight the need for caution; just because a tool is traditionally used it does not automatically follow that is an effective tool given it's reported limited sensitivity in recording smaller changes in joint range of motion [34]. Due to the overall weak agreement found for intra-rater and inter-rater reliability for both methods, it could be argued that a more reliable measurement tool should be utilised to adhere to evidence-based practice or that further research needs to be considered to further elucidate the effectiveness of goniometry in JPS assessment [4]. Although 2D IMC analysis may have associated initial costs and timing restraints, as highlighted by Smith et al [4], this method has demonstrated strong reliability for JPS assessment and although further more recent research into its reliability is warranted, this could potentially offer a more evidence-based alternative for clinical practice [19,35]. Currently, emerging evidence in relation to measurement of knee joint angles through smartphone applications could offer a cost-effective and easily clinically applicable alternative method of JPS measurement; however, further research is required to ascertain its reliability [25,26]. Such technology may be used in addition to audio bio-feedback, particularly at end of range measurements. This could enable repeatability training for the patient and a learning effect, particularly given the limitation in JPS measurements at extreme end range of motion as reported in these findings.

Conclusions

Overall, intra-rater and inter-rater agreement strength was weak and did not exceed substantial for either method. Generally AE and RE agreement was poor to moderate and greater error was reported for IMC for both assessors than goniometry. While using JPS is deemed an appropriate assessment of proprioceptive abilities, it is critical to be aware that by assessing proprioception in this manner, the error observed could have resulted from poor proprioception of the participants, measurement inaccuracy by the assessors or a combination of both factors and this cannot be ascertained. While these findings in isolation are insufficient to deem goniometry or IMC as unreliable measurement tools, they do have clinical implications, urge the use of caution and highlight the need for further research, particularly on the use of smartphone Apps for assessing JPS in varying clinical populations.

Acknowledgements and Declarations

Ethics Approval: The study was approved by the University of East Anglia's Faculty of Medicine and Health Science Research Ethics Committee (Ref: 2012/2013-14).

Funding: No funding was received to support this study.

Declaration of interest: None to declare.

Figure and Table Legends

Figure 1: Image to present the output and methods of the IMC methods of JPS assessment.

Table 1: Intra-rater reliability of goniometry compared to IMC methods for Assessor 1.

Table 2: Intra-rater reliability of goniometry compared to IMC for Assessor 2.

Table 3: Inter-rater reliability of goniometry compared to IMC.

Supplementary Table 1: Intra-rater reliability of goniometry method.

Supplementary Table 2: Inter-rater reliability of goniometry method.

Supplementary Table 3: Intra-reliability of IMC method.

Supplementary Table 4: Inter-rater reliability of IMC method.

References

[1] Lephart SM, Riemann, BL, Fu F. Introduction to the sensorimotor system. In: Lephart SM, Freddie F, editors. Proprioception and Neuromuscular Control in Joint Stability, USA: Human Kinetics; 2000: xv to xxiv.

[2] Deshpande N, Connelly DM, Culham EG, Costigan PA. Reliability and validity of ankle proprioceptive measures. Arch Phys Med Rehabil 2003; 84: 883-89.

[3] Stillman BC. Making sense of proprioception: the meaning of proprioception, kinaesthesia and related terms. Physiotherapy 2002; 88: 667-76.

[4] Smith TO, Davies L, Hing CB. A systematic review to determine the reliability of knee joint position sense assessment measures. Knee 2013; 20: 162-69.

[5] Grob KR, Kuster MS, Higgins SA, Lloyd DG, Yata H. Lack of correlation between different measurements of proprioception in the knee. J Bone Joint Surg Br 2003; 84: 614-18.

[6] Jerosch J, Prymka M. Proprioception and joint stability. Knee Surg Sports Traumatolo Arthrosc 1996; 4: 171-79.

[7] Moezy A, Olyaei G, Hadian M, Razi M, Faghihzadeh S. A comparative study of whole body vibration training and conventional training on knee proprioception and postural stability after anterior cruciate ligament reconstruction. Br J Sports Med 2008; 42: 373-78.

[8] Sahin N, Baskent A, Cakmak A, Salli A, Ugurlu H, Berker, E. Evaluation of knee proprioception and effects of proprioception exercise in patients with benign joint hypermobility syndrome. Rheumatol Int 2008; 28: 995-1000.

[9] Tsauo JY, Cheng PF, Yang RS. The effects of sensorimotor training on knee proprioception and function for patients with knee osteoarthritis: a preliminary report. Clin Rehabil 2008; 22: 448-57.

[10] Brosseau L, Tousignant M, Budd J, Chartier N, Duciaume L, Plamondon S, O'Sullivan JP, O'Donoghue S, Balmer S. Intratester and intertester reliability and criterion validity of the parallelogram and universal goniometers for active knee flexion in healthy subjects. Physiother Res Int 1997; 2: 150-66.

[11] Stillman BC. An investigation of the clinical assessment of joint position sense. [online] Available at: < https://minerva-access.unimelb.edu.au/bitstream/handle/11343/38786/65887_00000246_01_Stillman.pdf?sequence=1> (Accessed on: 05.02.2015).

[12] Marks R. Further evidence of impaired position sense in knee osteoarthritis. Phys Res Int 1996; 1: 127-36.

[13] Marks R, Quinney HA, Wessel J. Proprioceptive sensibility in women with normal and osteoarthritic knee joints. Clin Rheumatol 1993; 12: 170-75.

[14] Fleiss JL. Reliability of Measurement. In: Fleiss JL, editor. The Design and Analysis of Clinical Experiments, USA: John Wiley and Sons Inc; 1999.

[15] Piriyaprasarth P, Morris ME. Psychometric properties of measurement tools for quantifying knee joint position and movement: a systematic review. Knee 2007; 14: 2-8.

[16] Norkin, CC, White DJ. Measurement of Joint Motion. 2nd ed. USA: Allied Health, Jean-Francois Vilain; 1995.

[17] Field, A. Discovering statistics using IBM SPSS statistics, 4th Edn, London: SAGE Publications; 2013.

[18] Corrigan JP, Cashman WF, Brady MP. Proprioception in the cruciate deficient knee. J Bone Joint Surg Br; 1992; 74: 247-50.

[19] Stillman BC, McMeeken JM. The role of weightbearing in the clinical assessment of knee joint position sense. Aust J Phys 2001; 47: 247-53.

[20] Olsson L, Lund H, Henriksen M, Rogind H, Bliddal H, Danneskiold-Samsøe B. Test–retest reliability of a knee joint position sense measurement method in sitting and prone position. Adv Phys 2004; 6: 37-47.

[21] Baker V, Bennell K, Stillman B, Cowan S, Crossley K. Abnormal knee joint position sense in individuals with patellofemoral pain syndrome. J Orthop Res 2002; 20: 208-14.

[22] Rothstein JM, Miller PJ, Roettger RF. Goniometric reliability in a clinical setting: elbow and knee measurements. Phys Ther 1983; 63: 1611-615.

[23] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-74.

[24] Petrella RJ, Lattanzio PJ, Nelson MG. Effect of age and activity on knee joint proprioception. Am J Phys Med Rehabil 1997; 76: 235-41.

[25] Milanese S, Gordon S, Buettner P, Flavell C, Ruston S, Coe D, O'Sullivan W, McCormack S. Reliability and concurrent validity of knee angle measurement: Smart phone app versus universal goniometer used by experienced and novice clinicians. Man Ther 2014; 19:569-74.

[26] 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: 1038-43.

[27] Lee HM, Liau JJ, Cheng CK, Tan CM, Shih JT. Evaluation of shoulder proprioception following muscle fatigue. Clin Biomech 2003; 18: 843-47.

[28] Rozzi SL, Lephart SM, Gear WS, Fu FH. Knee joint laxity and neuromuscular characteristics of male and female soccer and basketball players. Am J Sports Med 1999; 27: 312-19.

[29] Dover G, Powers M. Reliability of joint position sense and force reproduction measures during internal and external rotation of the shoulder. J Athl Train 2003; 38: 304-10.

[30] Pincivero DM, Bachmeier B, Coelho AJ. The effects of joint angle and reliability on knee proprioception. Med Sci Sports Exerc 2001; 33: 1708-712.

[31] Escamilla RF. Knee biomechanics of the dynamic squat exercise. Med Sci Sports Exerc 2001; 33: 127-41.

[32] Barrack RL, Skinner HB, Buckley SL. Proprioception in the anterior cruciate deficient knee. Am Sports Med 1989; 17: 1-6.

[33] Smolensky MH, Hermida RC, Castriotta RJ, Portaluppi F. Role of sleep-wake cycle on blood pressure circadian rhythms and hypertension. Sleep Med 2007; 8: 668-80.

[34] Gajdosik RL, Bohannon RW. Clinical measurement of range of motion: review of goniometry emphasizing reliability and validity. Phys Ther 1987; 67: 1867-872.

[35] Gribble P, Hertel J, Denegar C, Buckley W. Reliability and validity of a 2-D video digitising system during a static and a dynamic task. J Sports Rehabil 2005; 14: 137-49.

Figure 1: Image to present the output and methods of the IMC methods of JPS assessment.

Table 1: Intra-rater reliability of goniometry compared to IMC methods for Assessor 1.

Session	Angle (º)	AE (º) Goniometry	AE (º) Image Capture	AE – ICC (95%)	AE *Agreement Strength	RE (º) Goniometry	RE (º) Image Capture	RE – ICC (95%)	RE *Agreement Strength
1	20	4.92	5.97	0.12 (0.00, 0.43)	Slight	4.69	2.64	0.24 (0.00, 0.52)	Fair
	40	4.03	5.36	0.00 (0.00, 0.30)	Poor	3.25	0.08	0.00 (0.00, 0.26)	Poor
	75	4.11	6.69	0.00 (0.00, 0.16)	Poor	-0.11	-5.58	0.23 (0.00, 0.50)	Fair
	100	3.64	9.31	0.30 (0.00, 0.58)	Fair	-2.58	-6.14	0.37 (0.07, 0.61)	Fair
2	20	3.81	6.33	0.07 (0.00, 0.37)	Slight	3.36	2.83	0.28 (0.00, 0.56)	Fair
	40	3.33	5.28	0.16 (0.00, 0.16)	Poor	1.83	-1.11	0.14 (0.00, 0.43)	Slight
	75	3.19	10.50	0.01 (0.00, 0.19)	Slight	-2.08	-10.33	0.14 (0.00, 0.41)	Slight
	100	3.50	10.64	0.01 (0.00, 0.20)	Slight	-0.89	-9.53	0.22 (0.00, 0.52)	Fair

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Table 2: Intra-rater reliability of goniometry compared to IMC for Assessor 2.

Session	Angle (º)	AE (º) Goniometry	AE (º) Image Capture	AE – ICC (95%)	AE *Agreement Strength	RE (º) Goniometry	RE (º) Image Capture	RE – ICC (95%)	RE *Agreement Strength
1	20	7.89	9.47	0.10 (0.00, 0.43)	Slight	7.72	8.97	0.12 (0.00, 0.43)	Slight
	40	8.64	7.58	0.09 (0.00, 0.41)	Slight	8.31	6.36	0.17 (0.00, 0.47)	Slight
	75	5.67	8.86	0.19 (0.00, 0.46)	Slight	-1.28	-4.92	0.39 (0.09, 0.63)	Fair
	100	6.22	7.56	0.46 (0.17, 0.68)	Moderate	-4.39	-5.78	0.56 (0.29, 0.74)	Moderate
2	20	6.64	8.17	0.05 (0.00, 0.36)	Slight	5.81	7.11	0.28 (0.00, 0.56)	Fair
	40	7.31	7.50	0.00 (0.00, 0.30)	Poor	6.917	3.17	0.13 (0.00, 0.42)	Slight
	75	4.56	9.39	0.04 (0.00, 0.29)	Slight	-1.50	-7.33	0.28 (0.00, 0.55)	Fair
	100	5.17	12.53	0.25 (0.00, 0.56)	Fair	-3.33	-11.08	0.38 (0.00, 0.69)	Fair

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Table 3: Inter-rater reliability of goniometry compared to IMC.

Session	Angle (º)	AE (º) Goniometry (º) Assessor 1	AE (º) Goniometry (º) Assessor 2	AE (º) ICC (95%)	AE (º) *AS	AE (º) Image Capture (º) Assessor 1	AE (º) Image Capture (º) Assessor 2	AE (º) ICC (95%)	AE (º) *AS	RE (º) Goniometry (º) Assessor 1	RE (º) Goniometry (º) Assessor 2	RE (º) ICC (95%)	RE (º) *AS	RE (º) Image Capture (º) Assessor 1	RE (º) Image Capture (º) Assessor 2	RE (º) ICC (95%)	RE (º) *AS
1	20	4.92	7.89	0.27 (0.00, 0.53)	Fair	5.97	9.47	0.11 (0.00, 0.39)	Slight	4.69	7.72	0.24 (0.00, 0.50)	Fair	2.64	8.97	0.19 (0.00, 0.45)	Slight
	40	4.03	8.64	0.64 (0.00, 0.31)	Subs	5.36	7.58	0.00 (0.00, 0.20)	Poor	3.25	8.31	0.09 (0.00, 0.35)	Slight	0.08	6.36	0.02 (0.21, 0.28)	Slight
	75	4.11	5.67	0.00 (0.00, 0.21)	Poor	6.69	8.86	0.04 (0.00, 0.35)	Slight	-0.11	-1.28	0.00 (0.00, 0.33)	Poor	-5.58	-4.92	0.30 (0.00, 0.57)	Fair
	100	3.64	6.22	0.01 (0.00, 0.30)	Slight	9.31	7.56	0.33 (0.01, 0.59)	Fair	-2.58	-4.39	0.17 (0.00, 0.46)	Slight	-6.14	-5.78	0.30 (0.00, 0.57)	Fair
2	20	3.81	6.64	0.06 (0.00, 0.34)	Slight	3.36	5.81	0.15 (0.00, 0.44)	Slight	6.33	8.17	0.03 (0.00, 0.35)	Slight	2.83	7.11	0.37 (0.06, 0.62)	Fair
	40	3.33	7.31	0.00 (0.00, 0.23)	Poor	1.83	6.92	0.09 (0.00, 0.35)	Slight	5.28	7.50	0.07 (0.00, 0.37)	Slight	-1.11	3.17	0.15 (0.00, 0.44)	Slight
	75	3.19	4.56	0.34 (0.04, 0.59)	Fair	-2.08	-1.50	0.46 (0.17, 0.69)	Mod	10.50	9.39	0.26 (0.00, 0.54)	Fair	-10.33	-7.33	0.22 (0.00, 0.50)	Fair
	100	3.50	5.17	0.00 (0.00, 0.19)	Poor	-0.89	-3.33	0.30 (0.00, 0.56)	Fair	10.64	12.53	0.47 (0.18, 0.69)	Mod	-9.53	-11.08	0.47 (0.18, 0.69)	Mod

AE - Absolute error; AS – Agreement Strength; ICC – intra-class correlation coefficient; Mod – Moderate; RE – relative error; Subs - Substantial.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Supplementary Table 1: Intra-rater reliability of goniometry method.

Assessor	Angle (º)	AE (º) Session 1	AE (º) Session 2	AE – ICC (95%)	AE *Agreement Strength	RE (º) Session 1	RE (º) Session 2	RE – ICC (95%)	RE *Agreement Strength
1	20	4.92	3.81	-0.11 (-0.42, 0.23)	Poor	4.69	3.36	-0.13 (-0.43, 0.21)	Poor
	40	4.03	3.33	0.17 (-0.15, 0.46)	Slight	3.25	1.83	0.18 (-0.16, 0.47)	Slight
	75	4.11	3.19	0.26 (-0.06, 0.54)	Fair	-0.11	-2.08	0.28 (-0.03, 0.55)	Fair
	100	3.64	3.50	0.15 (-0.19, 0.46)	Poor	-2.58	-0.89	0.22 (-0.11, 0.49)	Fair
2	20	7.89	6.64	0.42 (0.12, 0.67)	Moderate	7.72	5.81	0.29 (0.02, 0.56)	Fair
	40	8.64	7.31	0.09 (-0.24, 0.40)	Slight	8.31	6.92	0.24 (-0.90, 0.52)	Fair
	75	5.67	4.56	0.05 (-0.28, 0.36)	Slight	-1.28	-1.50	0.29 (-0.05, 0.56)	Fair
	100	6.22	5.17	0.38 (0.73, 0.62)	Fair	-4.39	-3.33	0.27 (-0.06, 0.54)	Fair

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Supplementary Table 2: Inter-rater reliability of goniometry method.

Session	Angle (°)	AE (°) Assessor 1	AE (°) Assessor 2	AE – ICC (95%)	AE *Agreement Strength	RE (°) Assessor 1	RE (°) Assessor 2	RE – ICC (95%)	RE *Agreement Strength
1	20	4.92	7.89	0.27 (-0.03, 0.53)	Fair	4.69	7.72	0.24 (-0.05, 0.50)	Fair
	40	4.03	8.64	0.64 (-0.15, 0.31)	Substantial	3.25	8.31	0.09 (-0.13, 0.35)	Slight
	75	4.11	5.67	-0.12 (-0.42, 0.21)	Poor	-0.11	-1.28	-0.00 (-0.33, 0.33)	Poor
	100	3.64	6.22	0.01 (-0.25, 0.30)	Slight	-2.58	-4.39	0.17 (-0.15, 0.46)	Slight
2	20	3.81	6.64	0.06 (-0.21, 0.34)	Slight	3.36	5.81	0.15 (-0.15, 0.44)	Slight
	40	3.33	7.31	-0.01 (-0.20, 0.23)	Poor	1.83	6.92	0.09 (-0.13, 0.35)	Slight
	75	3.19	4.56	0.34 (0.04, 0.59)	Fair	-2.08	-1.50	0.46 (0.17, 0.69)	Moderate
	100	3.50	5.17	-0.12 (-0.40, 0.19)	Poor	-0.89	-3.33	0.30 (0.00, 0.56)	Fair

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Supplementary Table 3: Intra-reliability of IMC method

Assessor	Angle (º)	AE (º) Session		AE – ICC (95%)	AE *Agreement Strength	RE (º) Session		RE – ICC (95%)	RE *Agreement Strength
		1	2			1	2		
1	20	5.97	6.33	0.06 (-0.28, 0.38)	Slight	2.64	2.83	0.07 (-0.27, 0.39)	Slight
	40	5.36	5.28	-0.08 (-0.41, 0.26)	Poor	0.08	-1.11	0.08 (-0.25, 0.40)	Slight
	75	6.69	10.50	0.22 (-0.06, 0.49)	Fair	-5.58	-10.33	0.27 (-0.0, 0.53)	Fair
	100	9.31	10.64	0.41 (0.09, 0.66)	Moderate	-6.14	-9.53	0.26 (-0.05, 0.53)	Fair
2	20	9.47	8.17	0.32 (-0.00, 0.58)	Fair	8.97	7.11	0.43 (0.13, 0.66)	Moderate
	40	7.58	7.50	0.11 (-0.26, 0.42)	Slight	6.36	3.17	0.24 (-0.07, 0.51)	Fair
	75	8.86	9.39	-0.24 (-0.54, 0.10)	Poor	-4.92	-7.33	0.11 (-0.23, 0.41)	Slight
	100	7.56	12.53	0.29 (-0.02, 0.55)	Fair	-5.78	-11.08	0.36 (0.05, 0.61)	Fair

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).

Supplementary Table 4: Inter-rater reliability of IMC method.

Session	Angle (º)	AE (º) Assessor 1	AE (º) Assessor 2	AE – ICC (95%)	AE *Agreement Strength	RE (º) Assessor 1	RE (º) Assessor 2	RE – ICC (95%)	RE *Agreement Strength
1	20	5.97	9.47	0.11 (-0.17, 0.39)	Slight	2.64	8.97	0.19 (-0.08, 0.45)	Slight
	40	5.36	7.58	-0.12 (-0.41, 0.20)	Poor	0.08	6.36	0.02 (0.21, 0.28)	Slight
	75	6.69	8.86	0.04 (-0.27, 0.35)	Slight	-5.58	-4.92	0.30 (-0.03, 0.57)	Fair
	100	9.31	7.56	0.33 (0.01, 0.59)	Fair	-6.14	-5.78	0.30 (-0.03, 0.57)	Fair
2	20	6.33	8.17	0.03 (-0.29, 0.35)	Slight	2.83	7.11	0.37 (0.06, 0.62)	Fair
	40	5.28	7.50	0.07 (-0.23, 0.37)	Slight	-1.11	3.17	0.15 (-0.14, 0.44)	Slight
	75	10.50	9.39	0.26 (-0.08, 0.54)	Fair	-10.33	-7.33	0.22 (-0.09, 0.50)	Fair
	100	10.64	12.53	0.47 (0.18, 0.69)	Moderate	-9.53	-11.08	0.47 (0.18, 0.69)	Moderate

AE - Absolute error; ICC – intra-class correlation coefficient; RE – relative error.
*Kappa statistic boundaries as outlined in Landis and Koch (1977).