Validity and Reliability of a Smartphone and Digital Inclinometer in Measuring the Lower Extremity Joints Range of Motion

Walaa S. Mohammad¹, ², Faten F. Elattar³, Walaa M. Elsais⁴, Salameh O. AlDajah⁴

Affiliations: ¹Cairo University, Faculty of Physical Therapy, Department of Biomechanics, Giza, Egypt, ²Majmaah University, College of Applied Medical Sciences, Department of Physical Therapy, Al-Majmaah, Saudi Arabia, ³Cairo University, Faculty of Physical Therapy, Supervisor of Orthopedic Outpatient Clinic, Giza, Egypt, ⁴Isra University, College of Allied Medical Sciences, Department of Physical Therapy, Amman, Jordan

Correspondence: W.S. Mohammad, Majmaah University, College of Applied Medical Sciences, Al-Majmaah, 11952, Saudi Arabia. Email: w.mohammad@mu.edu.sa

Abstract

In clinical settings, available valid and reliable tools are important components in evaluating the lower extremity range of motion. Although the digital inclinometer is highly reliable compared to the universal goniometer, its availability and high cost impede its extensive use. Nowadays, smartphone applications have become widely available to clinicians for assessing the joint range of motion. The present study aims to assess the validity and intra-rater reliability of the smartphone application “Clinometer” for measuring hip, knee, and ankle sagittal ranges of motion, using the digital inclinometer as the reference standard. Active hip, knee flexion and ankle dorsiflexion and plantarflexion range-of-motion measurements were recorded in 102 young, healthy female participants on two separate occasions using Clinometer and a digital inclinometer. Pearson’s correlation coefficients (r) were used to evaluate the smartphone application’s validity against the digital inclinometer. To assess the reliability of the Clinometer app, the intra-class correlation coefficient (ICC), standard error of measurement (SEM), and minimal detectable difference (MDD) were used. Clinometer displayed excellent validity when compared to the digital inclinometer for hip and knee movements (r>0.90), while ankle ROM displayed moderate validity (r = 0.52-0.57). Additionally, Clinometer demonstrated excellent reliability (ICC > 0.90) for hip and knee sagittal plane motion and moderate reliability for the ankle sagittal plane motion (ICC = 0.53–0.67). Clinometer is a portable, low-cost, valid, and reliable tool for assessing active hip and knee range of motions and can be easily incorporated into clinical settings; however, it cannot be used interchangeably for ankle measures.

Keywords: smartphone application; digital inclinometer; ROM; validity and reliability; lower limb joints


Introduction

In a clinical setting, physical therapists encounter many lower limb (LL) pathologies, such as osteoarthritis, muscle strain, musculoskeletal disorders, and overuse injuries, which can limit the hip, knee, and ankle joints range of motion (ROM) and hinder the daily living activities (Brosseau et al., 1997; Char-
A standard goniometer is a physical therapy assessment tool used to measure the range of motion (ROM) limitation, plan the appropriate treatment interventions, and track the treatment effectiveness (Behnoush et al., 2016; Brosseau et al., 1997). Although the universal goniometer (UG) has been extensively used due to its low cost, convenience, and reliability (Norkin & White, 2016), its intra-rater reliability appeared to be higher than its inter-rater reliability. Despite the good general reliability of the UG, this reliability tended to vary from one joint to another and from particular movement to another (Boone et al., 1978; Rothstein et al., 1983; Watkins et al., 1991). In addition, improper goniometer application, such as bony landmark positioning, detection, and maintaining the centre of the goniometer during measurement, can impact the universal goniometer’s data validity and reliability (Gajdosik & Bohannon, 1987). Measurement reliability is essential in the clinical setting because patients are often treated and re-evaluated, either by the same or by different physical therapists (Gogia et al., 1987). Currently, a digital inclinometer (DI) has been used because of its higher reliability in comparison with the UG for ROM assessment (Carey et al., 2010; Roach et al., 2013; Santos et al., 2012); however, the only drawback for the DI is its higher cost compared to the UG.

Recently, the American Physical Therapy Association (APTA) has encouraged the integration of technology into the practice of physical therapy, because of its impact on the profession and patients by improving the treatment outcomes and/or reducing the costs of treatment (Swisher & Hiller, 2010). Additionally, the APTA’s code of ethics, section 6B, states that “Physical therapists shall take responsibility for their professional development based on critical self-assessment and reflection on changes in physical therapist practice, education, health care delivery, and technology” (Swisher & Hiller, 2010).

Nowadays, smartphone technology provides opportunities for assessing and treating patients in clinical facilities and during the follow-up of the patient progress (Wojciechowski, 2011). To date, many studies have assessed the validity and reliability of the smartphone application against the UG for various joints motion, for instance, the spinal column, knee, and ankle (Alawna et al., 2019; Jones et al., 2014; Mitchell, Gutierrez et al., 2014). However, the validity and reliability of the smartphone application are limited compared to those of the DI in measuring sagittal plane ROM and is still of novel concern due to the importance of these joints in gait. Accordingly, the purposes of the present study were a) To determine the validity of the smartphone “Clinometer” application, using the DI as the reference standard, and b) To explore the intra-rater reliability of Clinometer for quantifying hip, knee, and ankle sagittal ROM.

## Methods

### Participants

A group of 102 young, healthy female participants (mean ± SD age: 20.71 ± 0.99 years, height: 158.08 ± 4.95 cm, mass: 57.16 ± 10.95 kg, body mass index (BMI): 22.86 ± 4.05 kg/m²) participated in this study. The results of a priori power analysis (G*Power software) revealed a need for 46 subjects or more to be sufficient to create a statistical power of 0.95, at the pre-set a level of 0.05 and with a large effect size of 0.5. Participants were included if they had no pain at the time of assessment with a clear history of LL musculoskeletal or neurological injuries within the 12 months prior to participation in the study.

Subjects were screened by completing a questionnaire. If they reported a history of musculoskeletal or neuromuscular disorders, rheumatologic condition, or previous surgeries to the LL, they were excluded from the study. Subjects with an abnormal end-feel other than soft tissue approximation for flexion, presence of pain or lack of full movement of LL joints (hip, knee flexion, ankle dorsiflexion, and plantar flexion) were also excluded from the study. In total, five subjects were excluded from participation in the study on these grounds. Subjects who fulfilled the inclusion criteria provided written informed consent prior to participation. Ethics approval for this study was attained from the Institutional Human Research Ethical Committee (Ethics approval no: P.TREC/012/002280).

### Instrumentation

The current study used validity and intra-rater reliability design to verify if a smartphone application could be considered a valid and reliable measuring tool for assessing the sagittal plane ROM of hip, knee, and ankle. The validity of the smartphone application was assessed compared to the Baseline DI instrument (Model 12-1057, Fabrication Enterprises, Inc., NY), which is a validated and reliable instrument for assessing the ROM.

The current study utilized a smartphone application known as “Clinometer” (CA) version 4.7. It is accessible on the Google Play store (https://play.google.com/store/apps) for free download and as a paid app (US $2.93) on the Apple iPhone store. A clinometer is a tool for assessing angles of slope (or tilt) using all sides of the device plus the camera. Clinometer uses three various units of measure: degree, per cent, and topo to quantify both inclines and declines. Additionally, it allows for two-way calibration on all sides: For every four sides and the bubble mode, calibration can be done in two procedures, which provides calibration on uneven surfaces. Each path can be calibrated separately to attain optimum precision. The smartphone used in the current study was an iPhone 7 Plus running iOS 4.7.

An expert physical therapist with fifteen years of clinical experience conducted all ROM measurements in the same laboratory on two separate sessions, one week apart. The tested limb was randomly selected, and only active ROM for hip, knee flexion, ankle dorsiflexion, and plantar flexion were measured. The order for ROM assessment (hip and knee flexion; ankle dorsiflexion and plantar flexion) and the measuring tool (DI and CA) were randomly performed (http://www.randomization.com/). For each movement direction, three trials were documented then averaged for the analysis purpose.

### Procedure

Hip flexion ROM was assessed while participants assumed supine lying position (Charlton et al., 2015). For each measurement of the hip joint, zero points for both DI and CA were set first; then, the measurement of hip flexion was performed. For both DI and CA, the instrument was placed over the thigh’s central point (a midpoint between the lateral femoral condyle and greater trochanter measured from a relaxed neutral hip position). The participant was then instructed to flex her hip maximally with a flexed knee position. The hip flexion measurement was recorded when the participant experienced a firm end-feel. The maximum hip flexion is kept for three seconds, during which the measurement was recorded using both CA and DI. For measuring maximum knee flex-
ion range, the participant was positioned in prone lying with an extended knee in neutral rotation then DI was placed just above the lateral malleolus to record the zero point of the knee. This followed by asking the participant to flex her knee maximally and to keep this position for three seconds. The same measurement procedures were repeated using the CA. For measuring the ankle dorsiflexion and plantarflexion ROM, the participant assumed a prone lying position with the knee bent 90° (Rabin et al., 2014). The DI or CA was positioned in the lateral side of the foot (aligned with the fifth metatarsus) while measuring ankle ROM (Gouveia et al., 2014). For all measurements, consistent oral instructions were offered to ensure appropriate performance during the measurement and to confirm reaching the available end-range.

**Statistical analysis**

Initially, the normality assumptions of parametric statistical tests were assessed based on the Shapiro-Wilk test, while the variance homogeneity was tested using Levene’s test. To assess the validity of the CA compared to the DI, Pearson’s correlation coefficients (r) were calculated to examine the strength of association between the ROM scores of the CA and DI for the tested joints in sagittal plane movements. The power of the correlation was established according to the following standard: r < 0.19 was considered as a very weak correlation, r = 0.2–0.39 as weak, r = 0.40–0.59 as moderate, r = 0.6–0.79 as strong, and r = 0.8–1 as very strong correlation (Campbell & Swinscow, 2011). The level of statistical significance was set at P < 0.05. The statistical analysis was conducted with the Statistical Package for the Social Sciences (SPSS) (version 25, IBM, NY, USA) and Excel (Microsoft Office Excel, 2016). The intra-class correlation coefficient (ICC) was calculated to assess test-retest inter-session reliability of active hip, knee, and ankle ROM.

The ICC reliability values were interpreted as poor reliability (< 0.50); moderate (0.50 to 0.75); good (0.75–0.90); and excellent (> 0.90) (Portney & Watkins, 2009). The standard error of measurement (SEM) and 95% confidence intervals (CI) were computed to quantify the amount of error linked to the measurement. Calculation of SEM was made using the formula: SEM = √Σ deviation^2 / degree of freedom (Bland & Altman, 1996). Also, the minimal detectable difference (MDD) was analysed to calculate the minimum threshold of measurement to give more confidence that differences between measurements were real and outside the error range. The MDD is calculated from the following formula: (SEM*1.96*√2) (Weir, 2005).

**Results**

The validity of the CA device against the DI, introduced as means and standard deviations, and ICC are provided in Table 1. The CA exhibited significant very strong correlations (r>0.90) with the DI for hip and knee flexion ROM measurements; however, the two devices displayed significant moderate correlations for ankle joint ROM measurements.

**Table 1. Validity of the Digital inclinometer and Smartphone application for active hip, knee and ankle ROM. Mean data presented as mean ± standard deviation**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Digital Inclinometer</th>
<th>Smartphone application</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (°)</td>
<td>Mean (°)</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>112.51 ± 9.28</td>
<td>112.67 ± 8.86</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>130.42 ± 8.99</td>
<td>129.48 ± 8.41</td>
</tr>
<tr>
<td>Ankle dorsiflexion</td>
<td>17.44 ± 2.18</td>
<td>17.74 ± 1.97</td>
</tr>
<tr>
<td>Ankle plantar flexion</td>
<td>34.54 ± 3.78</td>
<td>34.69 ± 3.54</td>
</tr>
</tbody>
</table>

Note. ICC: Intraclass correlation coefficient; CI: confidence interval; p: significance set at <0.05.

Intra-tester reliability of the CA presented as ICC (95% CI), SEM, and MDD values are provided in Table 2. The ICC values for hip and knee flexion ROM exhibited excellent (> 0.90) repeatability apart from ankle dorsiflexion and plantarflexion ROM; which reflected values in the moderate range repeatability. Furthermore, the SEM and MDD values can also be seen in Table 2, ranging from 1.45–4.29° and 4.01–11.89°, respectively.

**Table 2. Intra-tester reliability results of the smartphone application (CA) for active hip, knee and ankle ROM**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Smartphone application</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC (95%CI)</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>0.92 (0.87–0.96)</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.95 (0.92–0.97)</td>
</tr>
<tr>
<td>Ankle dorsiflexion</td>
<td>0.67 (0.50–0.77)</td>
</tr>
<tr>
<td>Ankle plantar flexion</td>
<td>0.53 (0.23–0.70)</td>
</tr>
</tbody>
</table>

Note. ICC: Intra-class correlation coefficient; CI: confidence interval; SEM: standard error of measurement; MDD: minimal detectable difference.

**Discussion**

Although smartphone applications are easily used in clinical practice, it is critical to explore the validity, reliability, and limitations before using them. Having such applications without investigating their validity and measurement variability means that they have no clinical value (Wellmon et al., 2016). To the best of our knowledge, this is the first study to assess the intra-tester reliability and validity of a smartphone application Clinometer for measuring active hip, knee, and ankle sagittal plane ROM. The CA demonstrated a very strong to a moderate association for LL measurements when compared to the DI. Furthermore, The CA demonstrated excellent reliability.
for hip and knee ROM, and moderate reliability for the ankle ROM. Consequently, the present study encourages the use of the smartphone application (i.e., Clinometer) as an alternative to the DI in assessing LL sagittal plane ROM among healthy female participants.

The present study found that the validity and intra-tester reliability of the hip and knee ROM was greater than that of the ankle joint; however, the subsequent SEM and SDD scores were greater in hip flexion ROM, which may have contributed to a relatively greater standard deviation seen in hip flexion ROM, indicating a variation in performance across individuals. This finding is in close agreement with the results of Charlton et al. (2015), who reported excellent association of the smartphone application against the three-dimensional (3D) motion analysis system, with good intra-tester reliability (ICC=0.86). However, no direct comparison can be conducted due to the different populations, the smartphone application used (Hip ROM Tester against the 3D-motion analysis system), and testing the passive flexion ROM of the hip in the previous study.

A lack of similar studies in literature prevents direct comparison with the outcomes of the present study. However, previous studies stated that smartphone applications could be a useful measuring tool to evaluate knee ROM. Many researchers have demonstrated a high agreement of passive (Milanese et al., 2014; Ockendon & Gilbert, 2012; Santos et al., 2012), or active knee flexion ROM (Hambly et al., 2012; Jones et al., 2014) in healthy adults or knee osteoarthritic patients or those after knee arthroplasty (Mehta et al., 2017), and postoperative knee surgery (Pereira et al., 2017) against UG. Interestingly, the current study found excellent intra-tester reliability for the CA in evaluating knee flexion ROM. Such reliability indicated that the knee ROM data were reproducible over the two different testing sessions, with small measurement errors and great confidence in the measured data. This is in close agreement with prior studies that assessed the reproducibility of the iPhone goniometer app for evaluating knee ROM in osteoarthritic knee patients or those after knee arthroplasty (Mehta et al., 2017). They reported excellent reliability of the iPhone goniometer app (ICC=0.97) in evaluating active knee flexion compared to UG.

Moreover, the current study demonstrated a higher ICC value than that reported by Derhon et al. (2017), who showed a good intra-tester correlation (ICC > 0.85) of passive knee flexion ROM in healthy women. The SEM and MDD values in the present study seemed to be extremely similar to those of Mehta et al. (2017), who reported smaller SEM (2.72°) and MDD (6.3°) in assessing knee flexion ROM. However, a direct comparison cannot be made due to the use of a different mobile application, population, and the reliability conducted against UG rather than DI in our study. Consequently, it can be concluded that the CA has greater accuracy in measuring knee ROM when compared to the DI in healthy females.

Unlike the hip and knee, the CA demonstrated moderate correlation and agreement for ankle ROM compared to the DI. Although, the placement on the individual for the two instruments is consistent and the ROM measured by CA was nearly similar to that by DI with similar variability in both measurement tools in our study. Since changes up to 4.01 and 9.07 may be attributable to measurement error in dorsiflexion and plantarflexion ROM, respectively, therapists may be incapable of monitoring and evaluating the therapy on the basis of minor differences. The source of error may be related to movement measured, the regional anatomy, the complexity of joint motion and the choice of testing position which depends on the purpose of the measurement (Rome & Cowieson, 1996). No study in the existing literature provided the validity of CA with DI in a non-weight bearing ROM position or for measuring plantarflexion ROM; however, previous literature showed excellent validity and reliability for smartphone app weight-bearing dorsiflexion ROM against DI (Vohralik et al., 2015; Williams et al., 2013) or UG (Alawna et al., 2019). This difference may have contributed to the greater stretch attained on soft tissue structures while the participant positioned in a weight-bearing lunge test that could influence the result.

Moreover, the results of our investigation are partially consistent with the outcomes of Vohralik et al. (2015). The SEM value for ankle dorsiflexion ROM of the current study was lower than the values of Vohralik’s data (2.68°). However, their intra-tester reliability for the active ankle dorsiflexion ROM was relatively higher (ICC=0.76) than in our study. The difference may be due to different phone applications (iHandy Level) or due to the different measurement techniques (weight-bearing lunge test) used in the previous study. In the present study, the lack of consistency and heightened error in ankle measurements may also be due to poor validity. Consequently, the measure that has poor validity and reliability is not dependable and should not be utilized to obtain clinical decisions (Bhattacherjee, 2012; Norkin & White, 2016).

Regardless of its novel findings, the present study has a number of limitations. Firstly, the sample enrolled in the study was limited to young females. Such a study group could potentially not be able to detect the true variability across the demographic continuum and, consequently, additional research is required on males and also more patients’ cohorts. Moreover, the validity and agreement of only sagittal plane ROM were evaluated in this study. Therefore, whether the same results would be achieved when assessing different plane motions is unclear. Future studies should be conducted LL ROM on different planes (e.g., hip frontal plane measurements). Lastly, the study was restricted to non-weight bearing ankle ROM. Therefore, additional studies are required to identify the difference in validity and reliability of smartphone application for ankle ROM in both conditions (weight-bearing versus non-weight-bearing positions).

Conclusion

The results of the present investigation provided an easy-to-use, fast, portable, valid, and reliable approach for evaluating joint ROM in the sagittal plane. The CA and DI can be used interchangeably with confidence, as “Clinometer” possesses very strong validity. This study showed that the “Clinometer” had excellent repeatability in assessing hip and knee ROM in young, healthy females. Although ankle dorsiflexion and plantarflexion ROM possess values in the moderate range of repeatability, the two instruments cannot be used interchangeably in the clinic for assessing the ankle ROM in a non-weight-bearing position. Future research should be directed to provide the required knowledge about the other plane movements and different measurement positions.

Acknowledgements

The authors would like to thank Deanship of Scientific Research at Majmaah University for supporting this work under Project Number No. R-2021-4.
References


