Quantification and Reliability of hip internal rotation and the FADIR test in supine position using a smartphone application in an asymptomatic population

Objectives: The purpose of this study was to quantify and to report the intra- and inter-rater reliability of the hip IR ROM supine with hip and knee in 90° of flexion and for the Flexion-Adduction-Internal rotation (FADIR) test. Hip internal rotation measure in lying supine position with hip and knee in 90° of flexion revealed information on hip impairments. To date no simple quantification approach has been presented in this positioning. Therefore, the FADIR test has not been quantified yet.

Methods: Twenty participants (24.0 ± 2.1 years, 10 women and 10 men) without lower limb or back pain were recruited. Three raters evaluated each participant during two testing sessions, one day apart. A built-in smartphone compass application (SCA) was used to obtain the hip IR ROM in both procedures.

Results: Mean IR supine were 51.7° (±9.7) and 62.6° (±11.4) for men and women, respectively. Concerning the FADIR test, mean values were 41.8° (±9.64) and 50.1° (±8.0) for men and women, respectively. The mean intra- and inter-rater reliability coefficients were 0.80 and 0.72 for the hip IR and 0.75 and 0.40 for the FADIR test. The SEMs ranged from 4.8 to 8.3° (MDD: 13.3 to 22.9°) for the hip IR and from 4.6 to 10.3° (MDD: 12.8 to 28.6°) for the FADIR test.

Conclusion: Overall, the SCA is adequate to quantify the hip IR in lying supine position. However, the poor to moderate inter-rater reliability in the FADIR test and the size of MDD values suggest to standardize the FADIR test.
1. Introduction

The hip internal rotation (IR) is recognized as relevant information in the assessment of symptomatic hips (1, 2). According to the literature, the main hip disorders that restrict the hip IR range of motion are the femoroacetabular impingement (FAI) and the hip osteoarthritis (OA) (1, 2, 3). The decrease of hip IR range of motion is strongly correlated to their severity and seems to appear long before pain (2, 3, 4, 5).

The hip IR can be evaluated with the subject in different positions such as seated, prone or lying supine (1, 6). In the seated position, the subject sits upright on the edge of the table with the hips and knees flexed to 90° (1). According to the literature, the use of digital inclinometer showed good to excellent intra-rater reliability (ICC: 0.82, 0.93) while photographic measurement showed moderate inter-rater reliability in this position (ICC: 0.77) (1, 6, 7).

In prone position, the hip and knee are placed at 0° and 90° of flexion, respectively (6). According to Kouyoumdjian, P. et al (2012), photographic measurement showed an inter-rater reliability of 0.83. In this protocol, two examiners were needed to perform the measurement.

Lastly, in lying supine, the hip was in neutral position with the knee placed at 90° of flexion by hanging over the edge of the exam table. The study from Charlton et al. (2015) showed high intra-rater reliability coefficient with a digital inclinometer in this position (ICC: 0.94).

Although these studies showed good intra and inter-rater reliability with various tools in different positions, none of them provide quantification of the hip IR ROM at 90° of flexion with subject lying supine. According to the literature, the latter presents the possibility to link the lost of hip internal ROM with hip impairments such as FAI and OA (2, 4).
While the loss of motion arises long before pain, this position must be quantified on an asymptomatic sample. This quantification must provide reliable measurement with simple and accessible tools to act as a good screening assessment across a large population of clinicians.

The quantification of the hip IR in lying supine, with the hip at 90° of flexion may be difficult to perform. The inclinometer, by relying on gravity, can’t be used to quantify the hip IR in the horizontal plane of motion related to the ground. The use of the long-arm goniometer showed good intra-rater reliability in this position (8). However, two evaluators are needed to perform the measurement with this tool. To date, no paper reported a simple protocol to quantify hip IR by only one evaluator.

Recently, smartphone applications have been used to quantify different ROMs such as cervical rotation (9). While compass application did not rely on gravity, it could present a good alternative to quantify the hip IR in the horizontal plane of motion related to the ground. Also, while adduction seems to increase the impact between femoral head-neck junction and the superior acetabulum rim in presence of FAI, it must be interesting to quantify the internal rotation in combination with hip adduction (10). This motion represents the Flexion-adduction-internal rotation test (FADIR) used in the assessment of symptomatic hips (11). Contrarily of the Flexion-abduction-external rotation test (FABER), also used in the assessment of hip pain, the FADIR test has not been quantified yet.

This study aims to report the intra- and inter-rater reliability using the intra-class correlation (ICC) of the smartphone compass application to quantify hip IR in two testing positions; supine with hip and knee in 90° of flexion and the FADIR test. The first specific objective is to compare the observed means of hip IR obtained between the two positions. The second specific objective is to compare the means of hip IR ROM between men and women within each testing positions.
2. Methods

2.1 Study design

The three raters that took part in this research were all kinesiology graduates. Prior to their recruitment, none of them were familiar with the assessment of hip IR ROM and FADIR test. To standardize the procedures across raters and ensure familiarity with the tests, a certified athletic trainer with eight years of practice demonstrated the protocol. Thereafter, each rater practiced the entire protocol at least three but up to five times to be at ease before starting the assessment. Three teams of raters were formed. Each team referred to two raters: one assessor who performed the tests manually and one note taker who compiled the data on a paper form. In this manner, raters A and B formed team #1, raters B and C team #2, and raters C and A formed team #3. Noted that a single assessor can perform the present protocol by writing down every result after each test. This procedure was selected to decrease the protocol duration.

Every team of two raters performed the assessment protocol on every participant in two sessions conducted one day apart. Participants were tested at the same time of the day in both sessions. They were asked to keep the same daily routine prior to each assessment session. Teams and tests orders were randomized for each participant but were kept the same between sessions one and two. The assessment protocol was composed of the hip IR in dorsal decubitus (hip in 90° of flexion) and of the FADIR test. The assessment lasted approximately 30 minutes.
2.2 Participants

Twenty participants were recruited from an academic setting (10 females, 10 males; [mean ± SD] age: 24.0 ± 0.9 years, height: 1.69 m ± 0.04 m; body weight: 68.3 ± 4.6 kg, BMI: 23.8 ± 1.3 kg/m²). They declared to be asymptomatic and never had hip pain in the past. Volunteers were not included if they reported hip, knee, low back or sacroiliac pain within the preceding three months, or if they had any history of hip or low back surgery. Volunteers who played soccer, hockey or football for at least one year while exceeding 5hr/wk. for any or 7hr/wk. when combined were not included. This last criterion was used to help decrease the risk of asymptomatic intra-articular problems (e.g. FAI, osteoarthritis, labral tears) in the sample (12). Also, asymptomatic subjects were recruited to minimize the impacts of pain on the reliability in the proposed measurement approach. The study was approved by the institutional ethic committees (CER-17-236-07.16) and in accordance with the revised version of the declaration of Helsinki. All participants gave their written informed consent prior to take part in this study.

2.3 Instruments

Two smartphones were used (iPhone 6s®, using iOS version 11.1.1). The built-in compass application of the first smartphone, which relies on the in-built magnetometer, was employed to note down the Starting and Ending positions of the tested leg in the horizontal plane related to the ground during the hip IR in supine position (see Figure 1, illustrations A.1 and A.2). This compass application also served the purpose of noting the two positions of the tested leg in an oblique plane during related to the ground the FADIR test (see Figure 1, illustrations B.1, B.2 and B.3). A custom case was made to hold the first smartphone and ensure a hand-free protocol for the compass application. This case was stuck on a 40 x 70 x 140.5 cm (height x width x length) sponge. Two
bands of self-gripping tape passed through the sponge and around the proximal and frontal aspect of the participant’s leg just under the tibial tuberosity (see Figure 1, illustration C.1.). The built-in inclinometer application from the second smartphone was used to record the maximal adduction angle obtained during the FADIR test.

2.4 Procedure of assessment session

2.4.1 Pre-testing procedure

The height and body weight of each participant were measured using a mechanical scale. Participants were then asked to lie supine on an examination table. Rater A marked the tibial tuberosity of the tested lower limb by using a black ink pen. Rater C then attached the smartphone around the tested leg just below this mark by using the custom case.

2.4.1. Hip internal rotation in lying supine position with the hip in 90° of flexion

First, the rater acting as the assessor brings the hip and knee at approximately 90 degrees of flexion. The same rater then holds the anterior aspect of the knee with one hand while the other hand is placed under the Achilles’ tendon (see Figure 1, illustrations B.1. and B.2.). This positioning was considered as the Starting position. The rater acting as note taker writes down the data observed at the Starting position. Then, the assessor brings the hip in internal rotation until maximum range of motion is reached. The maximum range of motion was considered as attained when a firm end-feel was detected without compensatory motion from the pelvis. The assessor holds this position for two seconds so that the note taker can write down the data observed at this Ending position indicated on the compass application. The internal rotation between the Starting
and Ending positions was termed $IR_{ROM}$. In clinical setting, the assessor can easily retain the starting position and write it down after the testing procedure.

2.4.2 Flexion-Adduction-Internal Rotation test

The FADIR test execution was slightly adapted from previous study (13). Usually performed in one movement, the FADIR was sequenced in three successive steps in order to standardize the testing protocol. First, the rater acting as the assessor brings the hip and knee to 90 degrees of flexion. Secondly, the same rater brings the hip in maximal adduction (see Figure 1, illustration C.1.). This maximal adduction angle was considered as attained just before the rater felt the ipsilateral side pelvis starting to lift off the examination table. The rater acting as note taker positioned the second smartphone at the mid-thigh level along its longitudinal axis and used the inclinometer application to write down the maximal hip adduction angle observed, termed $ADD_{MAX}$ (see Figure 1, illustrations C.1. et C.2.). The description above defines the Starting position. At that position, the note taker writes the data indicated on the compass. Lastly, while maintaining maximum adduction angle, the assessor brings the hip in internal rotation until the maximum is reached. The assessor holds this position for two seconds so that the note taker can write down the data observed at this Ending position indicated on the compass (see Figure 1, illustrations C.2. et C.3.). The internal rotation between the Starting and Ending positions was termed $FADIR_{ROM}$. 

2.5 Data analysis and statistical method

Data analysis was performed using SPSS v.24.0. The statistical threshold was set at p<0.05 (two-tailed). Intra-class correlation coefficients (ICC_{2,1}) and their 95% confidence interval were computed to obtain intra- and inter-rater reliability coefficients for the IR_{ROM} and FADIR_{ROM}. The ICCs were interpreted based on the classification of Portney and Watkins (15). Values under 0.50 are considered as poor reliability, between 0.50 and 0.75 as moderate and between 0.75 and 0.90 as good. Values over 0.90 are considered as excellent reliability. The standard error of measurement (SEM) was calculated to document the absolute error for each variable. The minimal detectable difference (MDD) was reported to provide the minimal difference that needs to be
obtained, in a 95% confidence interval, to ensure that a significant change occurred (16). A two-way repeated measures ANOVA was used to compare the amount of internal rotation for the men and women subgroup for the IR_{ROM} and FADIR_{ROM}. Simple effects test was performed to observe the amount of internal rotation for men and women subgroups within each test.

3. Results

The intra-rater reliability coefficients for each rater in each test are presented in table 1. The corresponding SEMs and MDDs are also presented in the same table. The intra-rater reliability coefficients reached between 0.90 and 0.70 across the three raters with a mean ICC of 0.80 for the IR_{ROM}. The mean SEM and MDD were 5.44 and 15.09 degrees, respectively.

The intra-rater reliability coefficient calculated from the FADIR_{ROM} reached between 0.79 and 0.67. The SEM and MDD calculated from these coefficients were 5.73 and 15.91 degrees, respectively (see table 1).

<table>
<thead>
<tr>
<th>Rater</th>
<th>IR_{ROM} ICC (95%CI)</th>
<th>SEM (°)</th>
<th>MDD (°)</th>
<th>FADIR_{ROM} ICC (95%CI)</th>
<th>SEM (°)</th>
<th>MDD (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.79 (0.45-0.92)</td>
<td>4.56</td>
<td>12.64</td>
<td>0.79 (0.54-0.91)</td>
<td>4.36</td>
<td>12.10</td>
</tr>
<tr>
<td>2</td>
<td>0.90 (0.77-0.96)</td>
<td>4.88</td>
<td>13.53</td>
<td>0.79 (0.56-0.91)</td>
<td>6.11</td>
<td>16.93</td>
</tr>
<tr>
<td>3</td>
<td>0.70 (0.39-0.87)</td>
<td>6.89</td>
<td>19.10</td>
<td>0.67 (0.35-0.85)</td>
<td>6.73</td>
<td>18.71</td>
</tr>
<tr>
<td>Mean</td>
<td>0.80 (0.68-0.91)</td>
<td>5.44</td>
<td>15.09</td>
<td>0.75 (0.67-0.83)</td>
<td>5.73</td>
<td>15.91</td>
</tr>
</tbody>
</table>

The inter-rater reliability coefficients for each test are presented in table 2. The mean ICC for the inter-rater reliability of the hip internal rotation is 0.72. The mean SEM value is 6.55 and the MDD value is 18.14 degrees. The mean inter-rater reliability coefficient for the FADIR_{ROM} is
0.40. The SEM and MDD values, calculated from this coefficient are 9.28 and 25.74 degrees, respectively.

### Table 2.

Inter-tester reliability for the IR$_{ROM}$ and FADIR$_{ROM}$ for the three pairs of raters (ICC$_{2,1}$)

<table>
<thead>
<tr>
<th></th>
<th>IR$_{ROM}$</th>
<th>FADIR$_{ROM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC (95%CI)</td>
<td>SEM</td>
</tr>
<tr>
<td>Rater 1-2</td>
<td>0.83 (0.63-0.93)</td>
<td>5.49</td>
</tr>
<tr>
<td>Rater 1-3</td>
<td>0.68 (0.27-0.87)</td>
<td>6.01</td>
</tr>
<tr>
<td>Rater 2-3</td>
<td>0.65 (0.32-0.84)</td>
<td>8.14</td>
</tr>
<tr>
<td>Mean</td>
<td>0.72 (0.61-0.83)</td>
<td>6.55</td>
</tr>
</tbody>
</table>

The mean range of motion of internal rotation for each test is presented in Table 3. The two-way repeated measure ANOVA showed that the IR$_{ROM}$, measured with the subject in dorsal decubitus with the hip in 90° of flexion with no adduction showed significantly greater internal rotation range of motion when compare to the FADIR test ($p < 0.001$).

### Table 3.

Range of motion measurement for the internal rotation and FADIR test

<table>
<thead>
<tr>
<th></th>
<th>IR$_{ROM}$</th>
<th>FADIR$_{ROM}$</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (±SD)</td>
<td>57.18° (±13.23) *</td>
<td>45.97° (±11.75)</td>
<td>$p &lt; 0.001$</td>
</tr>
</tbody>
</table>

The two-way factor ANOVA showed a greater amount of general (IR$_{ROM}$ and FADIR$_{ROM}$) internal rotation ($F = 5.09, p=0.04$) for the women subgroup when compared to the men subgroup.
The simple effects test showed no significant difference between women and men subgroup within each test (table 4).

**Table 4.**

Range of motion measurement for men and women in both tests

<table>
<thead>
<tr>
<th>Total ROM*</th>
<th>IR ROM</th>
<th>FADIR ROM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Mean (±SD)</td>
<td>46.78 (±10.71)</td>
<td>56.38 (±11.50)</td>
</tr>
</tbody>
</table>

*Significant difference between men and women (p = 0.04)
4. Discussion

The main objective of this study was to propose a new and simple quantification approach for the hip IR in lying supine position using a smartphone compass application. The smartphone was chosen due to its accessibility and simplicity of use (19). Therefore, this approach was adapted to quantify the FADIR test. Contrarily to the FABER test, the FADIR test has not been quantified yet.

Multiple studies have reported a relationship between a decrease in hip IR ROM in lying supine position and different hip impairments such as hip OA and FAI (2, 17, 18). As stated by Wyss et al. (2007), at 90° of flexion, the anterior femoral neck faces cephalic and the internal rotation brings the femoral neck closer to the superior acetabular rim. The authors from these studies reported that this assessment might act as a screening tool for early stages of FAI. Although previous studies have reported hip internal rotation ROM, none of them provide a quantification approach in this specific position (1, 6, 7).

The results from our study showed good to excellent intra-rater reliability for both procedures (hip IR and FADIR test) for two of the three raters involved. The weaker level of inter-rater reliability (from poor to moderate) suggests that the approach proposed might require a substantial standardization effort among raters. Despite the lower inter-rater reliability coefficients, the good to excellent intra-rater reliability coefficients showed that this new protocol is reliable to quantify hip internal rotation in lying supine position with hip and knee in 90° of flexion in a context of clinical follow-up done by the same evaluator.

The first specific objective was to compare the level of internal rotation in lying supine position and in the FADIR test. The internal rotation obtained was significantly greater in the lying supine position. The second specific objective was to compare the hip IR ROM for men and
women in both tests procedures. According to our results, the women subgroup had a significantly greater internal rotation ROM in both tests procedures when compared to men subgroup.

4.1 Hip IR in lying supine with hip and knee in 90° of flexion

The approach proposed for the quantification of hip IR ROM in lying supine position showed good to excellent intra-rater reliability for two of the three raters (max observed of 0.9 rater # 2). Previous studies reported intra and inter-rater reliability for the quantification of hip internal rotation in different positions but not in lying supine with the hip and knee flexed at 90°.

Studies from Krause et al. (2015) and Charlton et al. (2015) showed an intra-rater reliability of 0.84 in seated position while using a digital inclinometer and a smartphone, respectively. Charlton et al. (2015) also reported intra-rater reliability of 0.93 in lying supine position with no flexion at the hip using a smartphone. Nussbaumer et al. (2010) reported test-retest reliability of 0.90 while using a goniometer in supine position with hip and knee in 90° of flexion. Although the authors reported excellent reliability, two assessors were needed to perform the measurement and a belt was used to control the unmonitored movement from the pelvis. The approach presented in our study show similar intra-rater reliability coefficients when compare to previous studies (1,7,8).

Kouyoumandji et al. (2012) reported inter-rater reliability in three different positions such as supine, prone and seated. The inter-rater coefficients were respectively 0.80, 0.83 and 0.77. In this study, the authors used different tools such as a belt and cylindrical supports to control both the pelvis and the abduction and adduction motions at the hip. In our study, results illustrated lower inter-rater reliability for the quantification of hip IR ROM in lying supine position considering that good coefficient was noted for only one pair of raters (0.83 raters 1-2).
As cited above, the MDD values described the amount of internal rotation needed to ensure a significant change. The lowest MDDs reported in our study were 12.64° and 15.21° when related to the intra- and inter-rater reliability, respectively. As a comparison, the study from Krause et al. (2015) reported MDD values of 8.6 degrees in intra-rater setting. Contrarily to our study, Krause et al. (2015) used the seated position. This position might have helped to standardize the rotational axis of the lower limb. This standardization might have decreased variability across raters. It is known that the MDD values are highly dependent of the SEM calculation, also dependent of the amount of standard deviation (SD) (16). The differences between the SD of our study (±13.23°) and the one from Krause et al. (±7.7°), can explained why the MDD values are more important in our work.

The mean internal rotation ROM obtained in our study was 57.18 ±13.23°. This result is higher than the hip IR ROM reported in previous studies. The studies cited above reported hip IR ROM between 29.6 and 34.3 degrees (1, 6, 7, 8). It has been reported that the tension within the capsular ligaments is different between hip positioned in flexion compared to extension (21). The hip flexion position, releasing the capsular ligaments, might have brought the biggest difference when compared to previous studies.

Our results showed difference between gender for the hip IR ROM with the women subgroup having greater internal rotation ROM when compare to men. These results are supported by the study of Quatman, C. E. et al. (2008) showing a higher generalized joint laxity in women when compare to men for the same group of age.
4.2 FADIR test

This is the first study to quantify the hip internal rotation while performing the FADIR test. This new quantification showed good intra-rater reliability for two of the three raters. The inter-rater coefficients are lower showing more variability in the measurement.

The variability observed in both intra- and inter-rater reliability might come from different aspects. The maximal adduction positioning placed the hip in an oblique plan of motion bringing difficulties to perform the internal rotation while maintaining the same rotational axis. Therefore, the technique used to maintain the maximal adduction positioning might have differed across raters, affecting the FADIR\textsubscript{ROM}. These variabilities might explain the difference between intra- and inter-rater reliability for the FADIR\textsubscript{ROM}. The standard error of measurement and MDDs, calculated for the FADIR test were higher than those calculated for the hip IR\textsubscript{ROM}, showing greater variability.

As for all studies, our study presents some limitations. The first limitation is the absence of a symptomatic group in the sample. Although it could have showed differences between asymptomatic and symptomatic participants, this group would have brought new limitations. Indeed, the individual differences related to the manifestation of pain could have interfered in the reliability calculation. Therefore, it is difficult to clearly define the presence of FAI. Also, the type of FAI within men and women groups is different and might have brought inter-subject variability.

The absence of restrictive tools to maintain both, the pelvis and the lower limb is a choice to show the simplicity of the approach. However, further studies might use these tools paired with our approach and may see improvement in the reliability coefficients.
5. Conclusion

This is the first study to quantify the amount of hip internal rotation in lying supine position as generally carries out in clinical practice. The smartphone help to make an objective evaluation and its location was standardised using a precise anatomical landmark. While further steps are needed to validate the measurement with hip anatomical architecture, the simplicity of the approach allows the possible clinical transfer. Therefore, the absence of restrictive tools helps to adopt natural positioning for both, the clinician and patient. Lastly, to improve intra- and mostly inter-rater reliability coefficients, further studies should decrease the subjectivity of the hip internal rotation quantification.
References


