POSTPRINT VERSION. The final version is published here: St-Pierre, M. O., Boivin, K., Fontaine, N., Saadé, N., & Sobczak, S. (2021). Influence of Standardized Procedures on the Reliability of Hip Clinical Assessment. Journal of Manipulative and Physiological Therapeutics, 44(2), 137-145. https://doi.org/10.1016/j.jmpt.2020.09.003

Influence of Standardized Procedures on the Reliability of Hip Clinical Assessment

Objectives: This study evaluated a standardized and personalized approach to verify effects of conditions on intra- and inter-rater reliability, standard error of measurement (SEM) and minimal detectable difference (MDD) for provocative tests and range of motion (ROM) tests used in hip pain assessment: the flexion-adduction-internal rotation (FADIR), the flexion-abduction-external rotation-extension (FABER) test and the hip internal rotation in 90° of hip flexion (HIP IR).

Methods: 19 participants (23.79 ± 2.04 years, 10 women and 9 men) without lower limb or back pain were recruited. Three raters evaluated each participant during two testing sessions, one day apart. Raters performed the three tests in four conditions: classic (C), controlled pressure duration (CPD), subject-specific position (SSP) and mixed condition (M = CPD + SSP).

Results: For intra-rater reliability, the CPD condition showed the highest intra-class correlation (ICC Mean – [CI]) for hip IR_{ROM} (0.83 [0.53-0.94]) and $FADIR_{ROM}$ (0.75 [0.60-0.89]). The SSP condition showed the highest ICC for $FABER_{height}$ (0.71 [0.42-0.87]) and $FABER_{ROM}$ (0.62 [0.27-0.83]). Concerning inter-rater reliability, the classic condition presented the highest (ICC) for FABER variables (height: 0.54 [0.28-0.76], ROM: 0.58 [0.32-0.79]) and hip IR ROM (0.72 [0.51-0.87]). CPD condition showed the highest ICC for $FADIR_{ROM}$ (0.57 [0.32-0.78]).

Conclusion: For the conditions of this study, we found that CPD showed the highest ICC for hip IR_{ROM} and $FADIR_{ROM}$ and that SSP showed the highest ICC for $FABER_{height}$ and $FABER_{ROM}$.

INTRODUCTION

Assessment of hip conditions is composed normally of the patient history and physical examination. Although the patient history is clear, Van Trijffel E. et al. (2010) and Tijssen et al.(2012) have stated a large variability in the physical examination.

Clinicians, up to 98% of them, rely on hip ROMs such as the hip flexion and the hip internal rotation in 90° of flexion (3). According to Altman et al. (1991) and Krause et al. (2015), a decrease in hip internal rotation at 90° of flexion is associated with different hip impairments such as femoroacetabular impingement (FAI) or hip osteoarthritis.

Clinicians also use provocative tests when they suspect intra-articular problems such as FAI and labral lesions. The two most used tests are the flexion-adduction-internal rotation (FADIR) and the flexion-abduction-external rotation-extension (FABER) (3). However, Tijssen et al. (2012) described, in a literature review, an important uniformity in their executions. They stated that future researches should investigate the uniformity in the execution of these tests to improve both, the validity and the diagnostic accuracy.

Previous studies presented a quantitative approach for hip internal rotation range of motion (ROM) and FABER test (6, 7, 9). Although a large number of clinicians used the FADIR test, no quantification of ROM has been reported yet. The quantification of hip IR ROM has been performed in different subject positions such as seated, prone or supine, using different tools (5, 6, 7). From these studies, the intra-rater reliability was ranged between 0.84 and 0.94 and the inter-rater reliability between 0.80 and 0.93. None of these studies provides a simple quantification approach for the internal rotation in 90° of hip flexion.

The FABER test is useful to assess multi-directional ROMs (8). Therefore, limited ROM for FABER test might indicate hip pathology (8). Recently, Bagwell et al. (2016) quantified the

FABER height in an asymptomatic population using both, a digital inclinometer and a carpenter square. They reported intra-rater reliability of 0.76, 0.84 and 0.86 across the three raters, respectively. The inter-rater reliability was 0.68. The FABER height describes the distance between the lateral aspect of the knee and the examination table.

As stated previously, hip IR ROM, FABER and FADIR tests are used when clinicians suspect intra-articular hip impairments (FAI or hip osteoarthritis). The FAI affects the young and athletic population by modifying the osseous architecture due to repetitive micro-trauma (10, 11, 12). In a two years follow-up study, the prevalence of cam-type deformity increased from 2 to 18% in young soccer players (11-13 years old) (10). Lahner et al. (2014) showed that 62.5% of semi-professional soccer players had signs of FAI compared to only 27% in the amateur players group. Athletes use their hips in a specific manner and might create osseous bumps in a specific area of their femoral head and neck junctions. Following this assumption, we though that using a subject-specific hip position combining the hip flexion and abduction might helped to improve the specificity of these tests in detecting hip FAI. Previous studies on FAI note the link between ice hockey and hip osseous modifications (8, 13, 21). In concordance with these studies, we designed a subject-specific motion based on a skating movement. To date, no study has used a subject-specific position (specific to the sporting activity) to assess hip IR ROM, FABER or FADIR test.

The aim of this study was to present a new subject-specific (personalized) and standardized approach of quantification for hip internal rotation, FADIR and FABER tests and to note how subject-specific angle and the standardization influenced the intra- and inter-rater reliability of these tests.

METHODS

This reliability protocol was assessed on an asymptomatic population by three raters in order to validate the measurement procedures proposed. The raters came from a kinesiology background and they were unfamiliar with the three tests cited above. Lastly, this study was conducted to create a simple protocol that could be used by raters from different backgrounds and experiences in clinical assessment.

Participants

Nineteen participants, reporting no hip pain in their past life, were recruited from an academic student population (10 females, 9 males; (mean± SD) Age: 23.8 ± 2 years, Height: 1.69 ± 0.1m; Weight: 67.7 ± 10.4 kg and BMI: 23.8 ± 3.1 kg/m^2). Volunteers who played soccer, hockey or American football for at least one year while exceeding 5 hr./wk. for any of these sports or 7 hr./wk. when combined, were excluded to decrease the risk of asymptomatic intra-articular problems such as FAI, labral tears or osteoarthritis (13).

Ethics

The study was approved by the Ethics Subcommittee of the Anatomy Teaching and Reasearch Laboraty (SCÉLERA) of the Université du Québec à Trois-Rivières (SCELERA_19_05) and in accordance with the revised version of the declaration of Helsinki. All participants gave their written informed consent prior to participation in this study.

Study design

Three kinesiologists took part in this study (Rater A, B, C). To standardize the testing procedures across raters and ensure a proper familiarization, a certified athletic trainer with eight years of experience demonstrated each test. Thereafter, each rater practiced the entire protocol between three and five times with and without each instrument (see instrumentation section). Three teams of two raters were formed (A-B, B-C, C-A). The first rater referred to an ‘assessor, and the second to a ‘note taker. Assessors performed the tests manually and note takers compiled the data on a form. Forming teams helped to decrease the duration of each assessment and limited error related to data collection avoiding assessors to take notes. Teams, tests orders and tests conditions were randomized for each participant but assessors and note takers kept their function across the two assessment sessions. Participants were tested at the same time of the day and were asked to keep the same daily routine prior to each assessment to limit within-subject variability. The assessment protocol lasted between 45 to 55 minutes.

Instrumentation

Two smartphones were used to perform the assessment. The first smartphone, with a built-in inclinometer, was used to record the FABER ROM in degrees (Figure 2). The FABER height, measured between the lateral aspect of the knee and the examination table, was obtained using a carpentry scale (Figure 2). The same smartphone was used to record the internal rotation of HIP IR and FADIR test using the built-in compass application (Figure 3A-B). Based on previous study, the intra-rater ICC and validity for the iPhone® compass application were respectively 0.96 − 0.98; 95% CI [0.93–0.99] and r = 0.835 (Pearson’s correlation coefficient) $p < 0.001$ when compared to an universal goniometer (SAEHAN Grip ™ Rulong 20 cm, Belgium) (14).

The second smartphone was solely used for the FADIR test to obtain the maximal adduction angle. The dynamometer (Hand-held dynamometer, Model: 01163, Lafayette Industries, USA) was used to control pressure duration to five seconds (Figure 3-C1). The dynamometer model used in this study allowed us to set an audible alarm after five seconds of pressure. Therefore, when the raters started pressing on the lower limb, the timer started, and the alarm sounded to indicate when the application of pressure was completed. Lastly, a long arm goniometer was used to position the lower limb according to the subject-specific angle (flexion or abduction).

Assessment procedure

Pre-testing procedure

Height and weight of each participant were measured using a mechanical scale. Rater A marked the popliteal fossa, the femoral lateral epicondyle, the tibia tuberosity of the tested leg and both posterior-superior iliac spines. Those anatomical landmarks help to obtain the subject-specific abduction and flexion angles (Figure 1). The tibial tuberosity represented the smartphone location and was kept across all participants, tests and sessions.

The subject-specific movement

Participants were asked to perform two subject-specific movements. Movements were hip flexion and abduction motion (based on a skating movement) in a standing position. Hip flexion position was used as an alternative condition to the 90° of flexion for IR_{ROM} and $FADIR_{ROM}$ tests. Sub-maximal abduction motion was used to position hip for FABER variables (height and ROM). These ROMs were used in the *Subject Specific Position condition* (SSP) and *mixed condition* (M).

The rater A explained, demonstrated and measured both subjects-specific angle (flexion and abduction). Before performing movements, participants were asked to bend over and bend knees while keeping feet at the width of the pelvis. Participants were asked to observe a mark on the wall to keep the head in a straight line with the torso and keep the pelvis parallel to the floor throughout the movement. This represent the starting position.

The subject-specific hip abduction was measured on participants' pictures performing the abduction movement (Figure 1. A). The subject-specific hip flexion was measured using pictures taken in the sagittal plane with participants maintaining the starting position (Figure 1.B.).

The first component of the angle measurement task represented the intra-rater reliability of the measurement taken at the beginning and at the end of the first session (45 minutes between both measurements). The angle measurement was taken from the same picture.

The second component represented the inter-session reliability. This reliability was measured from two movements and two different pictures. Each movement-picture pair were assessed at the beginning of each session. This entire process described the participant movement and position variabilities and the variability from the rater performing the angle measurement. These two components were evaluated to calculate the measurement task reliability.

After both movements, rater A measured angles (flexion and abduction) with the iPad application (Hudl technique, Version: 6.1.0) (Figure 1). The iPad was set at the same height and angle using a tripod (AceTaken Mount®,). The subject-specific abduction was obtained using the largest ROM. Abduction angle was measured between a horizontal line passing by the two posterior-superior iliac spine (PSIS) and a vertical line passing by the PSIS and the ipsilateral fossae popliteal (Figure 1. A) on each side. The subject-specific flexion angle was calculated between a line passing by the ear and the greater trochanter and a line passing by the greater trochanter and the knee lateral epicondyle (Figure 1. B).

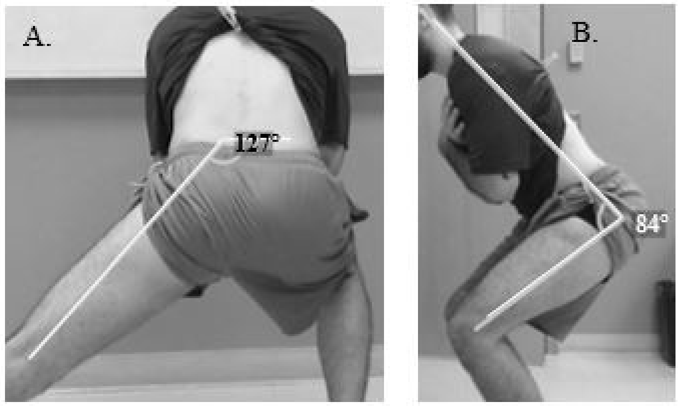

Figure 1. *A.* Abduction angle was measured between a horizontal line passing by the two posteriorsuperior iliac spine (PSIS) and a vertical line passing by the PSIS and the ipsilateral fossae popliteal on each side. The participant was asked to reach a maximal abduction motion while keeping stability. *B.* The subject-specific flexion angle was calculated between a line passing by the ear and the greater trochanter and a line passing by the greater trochanter and the knee lateral epicondyle. The subject-specific hip flexion was measured in a static and comfortable position. Both photos were taken using the Hudl application.

Testing procedure

Conditions

Each test was performed under four different conditions. The first condition referred to the Classical Condition (C) (described below). The second condition is characterized by the controlled pressure duration (CPD) using the dynamometer (Figure 3-C1). The third condition is characterized by the subject-specific position according to the skating movement (Abduction and Flexion) (Figure 1) (SSP). The fourth condition regroups the CPD and the SSP described as the mixed condition (M).

The FABER test

Concerning the C condition, the FABER test positioning was the same as described in previous studies (15, 2, 9). Participants were asked to lie in supine position on the examination table. Firstly, the assessor brought the external malleolus over the contralateral distal aspect of the thigh (Figure 2). At this moment, the assessor used the long-arm goniometer to measure the hip abduction angle. The fixed branch was placed parallel to the line between both anterior-superior iliac spines. The rotational axis was placed over the hip rotational axis. The mobile branch follows the centerline of the thigh. Concerning the SSP and M conditions, the assessor placed the tested lower limb in a position to obtain the hip abduction previously obtain (hip abduction motion). One hand stabilized the contralateral anterior-superior iliac spine (ASIS) of the tested lower limb. The other hand was placed, without any pressure, over the lateral aspect of the knee of the tested lower limb. This position represented the *starting position.* The note taker documented the vertical distance (table from tibial tuberosity) using the carpentry square. Right after, the note taker documented the inclination of the lower limb using the built-in digital inclinometer on the smartphone. When ready, assessors put a downward overpressure on the knee of the tested lower limb until the maximum ROM was reached. Concerning the CPD condition, raters pushed on the medial aspect of the knee with the dynamometer while respecting its application time. The maximum ROM was described as attained when the contralateral ASIS started to push in the assessor's hand. When the assessor felt this end-point position (or the five seconds alarm for the CPD and M conditions), the note taker documented the vertical distance and angle and in the same order as described previously. The vertical distance was termed as $FABER_{height}$ and the ROM was termed as $FABER_{ROM}$.

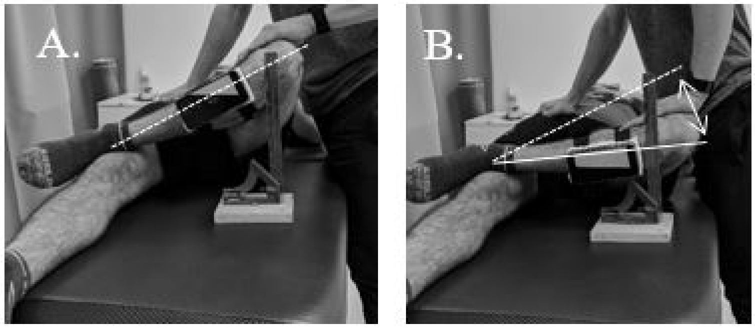

Figure 2. A. The starting position with the patient lying supine. B. The ending position for the FABER test. Both picture views in a transversal plan (Angle A- Angle B = $FABER_{ROM}$ (inclinometer), Height A – Height B = $FABER_{Height}$ (carpenter scale)).

Hip internal rotation ROM

Concerning the C condition, hip internal rotation ROM was performed with participants in supine position with hip and knee at 90° of flexion. First, the rater acting as the assessor brought both, the hip and knee to 90 degrees of flexion. The same rater then held the anterior aspect of the knee with one hand while the other hand was placed under the Achilles' tendon.

At this moment, the note taker used the goniometer to replicate the hip flexion angle for the SSP and M conditions. The rotational axis of the goniometer was positioned on the greater trochanter. The fixed branch was positioned to create a line between the greater trochanter and the participant's axilla. The mobile branch creates a line between the center of the greater trochanter and the femoral lateral epicondyle. The assessor placed the tested lower limb in a position to obtain hip flexion previously obtain (hip flexion motion). This position was noted as the *Starting position*. The rater, acting as the note taker, documented the data observed at this position.

Then, the assessor brought the hip in internal rotation until the maximum range of motion was reached. Concerning the CPD condition, raters pulled on the medial and distal aspect of the leg with the dynamometer while respecting its application time. The maximum range of motion was considered as attained when a firm end-feel was detected without compensatory motion from the pelvis as described by Prather et al. (2010). Assessor held this position for two seconds and the note taker documented the data observed at this Ending position indicated on the compass application. The internal rotation between the Starting and Ending position was termed as IR_{ROM}.

The FADIR test

Concerning the C condition, the FADIR test procedure was slightly adapted from previous studies (2, 16). Usually, the FADIR test is performed in one movement. However, to standardise the testing protocol, it was sequenced by three successive steps. First, the assessor flexed the hip and knee to 90 degrees. In this position, concerning both SSP and M protocol, the note taker used the goniometer to guide the assessor for the subject-specific hip flexion with the same protocol described above (hip IR ROM). While keeping the subject-specific flexion, the assessor placed the hip in maximal adduction. This ADD_{MAX} position was obtained just before the rater felt the ipsilateral side pelvis starting to lift off the examination table. When reached, the note taker placed the second smartphone at the mid-thigh level along its longitudinal axis and used the inclinometer application to measure the maximal adduction. As the description above referred to the Starting position, the note taker wrote the value indicated on the compass. Lastly, while maintaining maximum adduction and subject-specific hip flexion, the assessor placed the hip in internal rotation until the maximum ROM. Concerning the CPD condition, the raters pulled on the medial and distal aspect of the leg with the dynamometer while respecting its application time. The assessor held this position for two seconds and the note taker was able to read the value observed on the compass application at the ending position. The internal rotation between the Starting and Ending position was defined as ROM_{FADIR}.

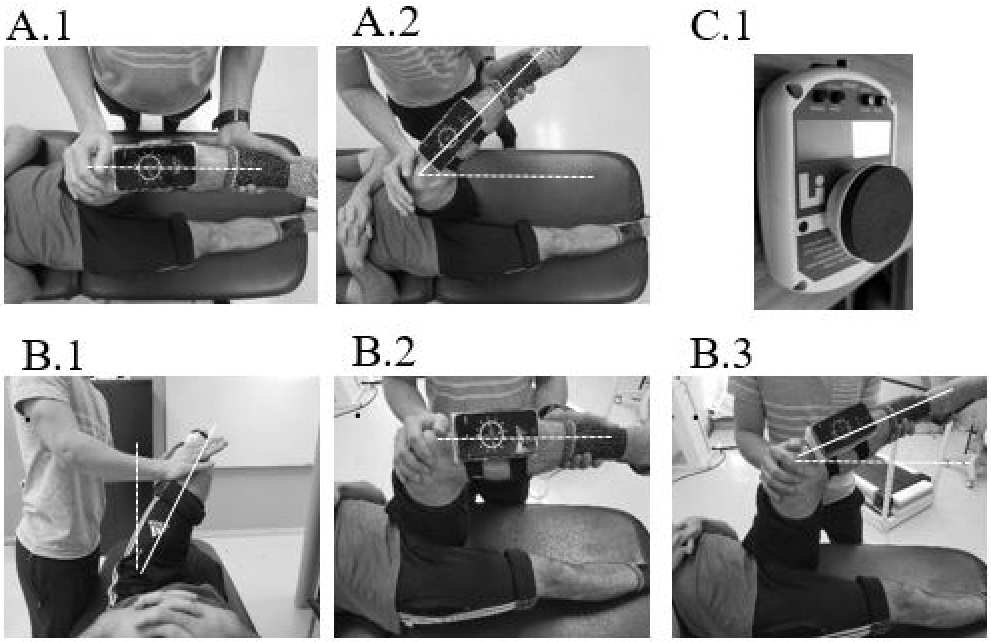

Figure 3. A1. Starting position with the patient lying in supine position with hip and knee in 90° of flexion. A1. Ending position of the hip internal rotation. B1. Maximal adduction position during the FADIR test. B2. Starting position of the FADIR test with the participant lying supine with hip in maximal adduction and 90° of flexion and the knee in 90° flexion. B3. Ending position of the FADIR test. C1. The algometer

<u>Data analysis and statistical method</u>

Data analysis were performed using SPSS v.24.0 and Statistica v.13.2. The statistical threshold was set at $p<0.05$ (two-tailed). The intra-class correlation coefficient (ICC 2,1), the standard error of measurement (SEM) and the minimal detectable difference (MDD) were computed for the rater angle measurement technique (intra-session) and for the complete process of the personalized angle measurement (inter-session). Intra-class correlation coefficients (ICC 2,1), SEM and MDD were also computed to obtain intra- and inter-rater reliability for these four variables: FABER ROM ($FABER_{ROM}$) and FABER height ($FABER_{height}$), Internal rotation ROM (IR_{ROM}) and FADIR ROM ($FADIR_{ROM}$). The ICC 2,1 was chosen according to the fact that raters represented a various population of raters and making the generalization of the results possible.

A two-way ANOVA models was used to compare the observed means ($IR_{ROM,}$ $FABER_{ROM}$, $FABER_{height}$ and $FADIR_{ROM}$) between the four conditions for each test (session, conditions). The post-hoc Tukey tests were used to note which condition differed significantly.

Results

The intra-session and inter-session angle measurement technique

The intra-session measurement technique (Rater A) and their related SEM and MDD for the abduction and flexion positioning are presented in Table 1. Intra-session reliability values were 0.80 and 0.95 for abduction and flexion positioning, respectively (Table 1). The inter-session reliability concerned both participants' positions and raters' variabilities. The abduction and flexion measurement reliability were 0.52 and 0.79, respectively (Table 1). The ROMs were measured with participants in a standing position.

Table 1. Intra- and inter-session reliability for the angle measurement technique

	Intra-session				Inter-sessions			
Positions	**Means (°)**	**ICC**	**SEM (°)**	**MDD (°)**	**Means (°)**	**ICC**	**SEM (°)**	**MDD (°)**
Abduction	123 ± 3	**0.80** (0.39-0.95)	1.5	4.1	123 ± 4	**0.52** (0.12-0.78)	3.1	8.5
Flexion	89 ± 3	**0.95** (0.81-0.99)	0.8	2.2	92 ± 12	**0.79** (0.53-0.91)	5.5	15.2

Test-retest: The ICCs were obtained from rereading the same picture at the beginning and ending of first session.

Inter-session: The ICCs were obtained with measures taken at the beginning of session one and two (complete process). ICC: intra-class correlation coefficient, SEM: standard error of measurement, MDD: minimal detectable difference.

The controlled pressure duration and the subject-specific position effects

The mean intra-rater reliability coefficients (ICC) for all variables and their corresponding SEMs and MDDs are presented in Table 2. The mean pressure (kg) are reported for the CPD and M due to the used of the dynamometer in these two conditions. The inter-rater reliability coefficients for all variables and conditions are presented in Table 3.

The smallest SEM and MDD values were not obtained in the same testing condition for every test. Concerning the $FABER_{height}$ the smallest SEM and MDD were obtained for the classic condition (C). Concerning the $FABER_{ROM,}$ the lowest SEM and MDD values were calculated for the subject-specific position (SSP). Lastly, concerning the hip IR_{ROM} and $FADIR_{ROM}$, the smallest SEM and MDD values were obtained while using the dynamometer only (CPD).

Table 2. Intra-rater correlation coefficients for each variable					
FABER$_{height}$					
Conditions	**Mean** ± SD **(cm)**	**ICC** (CI)	**SEM** ± SD (cm)	**MDD** ± SD (cm)	**Mean Pressure (kg)**
C	**13.2** ± 2.0	**0.60** (0.45-0.83)	**1.5** ± 0.3	**4.0** ± 0.8	NA
CPD	**13.5** ± 2.5	**0.61** (0.25-0.83)	**1.8** ± 0.2	**5.0** ± 0.6	13.66
SSP	**13.2** ± 2.5	**0.71** (0.42-0.87)	**1.6** ± 0.2	**4.3** ± 0.7	NA
M	**13.6** ± 2.0	**0.62** (0.26-0.83)	**2.1** ± 0.5	**5.7** ± 1.2	13.59
FABER$_{ROM}$					
Conditions	**Mean** ± SD **(°)**	**ICC** (CI)	**SEM** ± SD (°)	**MDD** ± SD (°)	**Mean Pressure (kg)**
C	**23.0** ± 4.2	**0.51** (0.14-0.76)	**3.5** ± 0.6	**9.6** ± 1.6	NA
CPD	**23.0** ± 4.7	**0.59** (0.22-0.81)	**3.7** ± 0.8	**10.2** ± 2.3	13.66
SSP	**19.7** ± 3.7	**0.62** (0.27-0.83)	**2.9** ± 0.2	**8.0** ± 0.6	NA
M	**20.0** ± 3.9	**0.46** (0.06-0.74)	**3.7** ± 0.9	**10.1** ± 2.4	13.59
IR$_{ROM}$					
Conditions	**Mean** ± SD **(°)**	**ICC** (CI)	**SEM** ± SD (°)	**MDD** ± SD (°)	**Mean Pressure (kg)**
C	**57.4** ± 12.2	**0.80** (0.54-0.92)	**5.7** ± 1.1	**15.9** ± 3.1	NA
CPD	**56.2** ± 11.5	**0.83** (0.53-0.94)	**5.1** ± 1.1	**14.1** ± 3.1	8.17
SSP	**57.9** ± 10.8	**0.77** (0.51-0.91)	**5.2** ± 1.8	**14.5** ± 5.1	NA
M	**54.3** ± 12.2	**0.72** (0.43-0.88)	**5.7** ± 1.0	**15.9** ± 2.7	8.11
FADIR$_{ROM}$					
Conditions	**Mean** ± SD **(°)**	**ICC** (CI)	**SEM** ± SD (°)	**MDD** ± SD (°)	**Mean Pressure (kg)**
C	**45.9** ± 10.0	**0.76** (0.48-0.89)	**5.8** ± 1.1	**16.2** ± 3.0	NA
CPD	**44.8** ± 10.5	**0.75** (0.60-0.89)	**5.6** ± 1.2	**15.6** ± 3.3	8.33
SSP	**45.1** ± 10.6	**0.74** (0.44-0.89)	**6.5** ± 0.9	**18.0** ± 2.4	NA
M	**44.6** ± 10.1	**0.71** (0.39-0.88)	**6.3** ± 0.9	**17.3** ± 2.4	8.34

The Table 3 shows the inter-rater reliability for each test. Three out of four variables show the higher ICC in the classic condition (FABER$_{height}$, FABER$_{ROM}$, IR$_{ROM}$). The highest inter-rater ICC for the FADIR$_{ROM}$ was obtained in the controlled pressure duration condition (using the dynamometer).

Table 3. Inter-rater correlation coefficients for each variable

Conditions	**FABER$_{height}$**	**FABER$_{ROM}$**	**IR$_{ROM}$**	**FADIR$_{ROM}$**
CC	**0.58** (0.32-0.79)	**0.54** (0.28-0.76)	**0.72** (0.51-0.87)	**0.40** (0.12-0.67)
CPD	**0.47** (0.19-0.72)	**0.38** (0.10-0.65)	**0.64** (0.41-0.82)	**0.57** (0.32-0.78)
SSP	**0.46** (0.19-0.70)	**0.33** (0.08-0.61)	**0.66** (0.43-0.83)	**0.39** (0.11-0.66)
M	**0.54** (0.54-0.91)	**0.50** (0.24-0.73)	**0.63** (0.31-0.83)	**0.48** (0.43-0.89)

The two-factor ANOVA revealed different F values. According to the F table for p-values, the only significant difference occurred in the $FABER_{ROM}$ conditions ($F_{3, 54} = 11.97$, $p < 0.001$, effect size: 0.40). The post-hoc Tukey test demonstrated that the subject-specific hip positions bring differences in the observed means between the following conditions: C vs SSP p = 0.009, C vs M p = 0.023, CPD vs SSP p = 0.01 and CPD vs M 0.026. The different conditions did not influence observed means of the FABER height, hip IR ROM and FADIR ROM test.

Discussion

As far as we are aware, this is the first study to control the pressure duration in the assessment of the hip IR, FABER and FADIR tests and to use a subject-specific position to assess those tests. The CPD and the SSP were used to verify their effects on reliability. Lowering the variability in the participant position (SSP) or the pressure duration (CPD) might improve reliability, SEMs and MDDs in the assessment of hip ROMs. It would be interesting for clinicians to find a balance between the classic test and the used of tools to improve reliability. As an example, clinicians could focus on test procedures such as the patient position, the clinician position, the pressure site and tools to restrict unmonitored motion from proximal joints. We hypothesize that generalized explanation (standardization) could improve the uniformity of testing and by this way improve the follow-up of patients with hip pain.

Several studies reported the significance of hip physical evaluation to enhance the assessment of various hip impairments such as femoroacetabular impingement (FAI) or hip osteoarthritis (2, 3, 15, 17 23). Most of these hip clinical assessments consist of the hip ROMs and the use of pain provocative tests (3).

As stated by Martin H.D. et al. (2010), 70% and 50% of the clinicians use the FADIR and FABER tests, respectively. However, as reported by Van Trijffel E. et al. (2010) and Tijssen et al. (2012) there is high variability in the executions of hip provocative tests. The hip IR ROM is largely used in different positions (3, 6). Although a decrease in hip IR_{ROM} at 90° of flexion is linked with FAI or hip osteoarthritis, no studies provide a simple quantification of this ROM.

Intra- and inter-rater reliability

ICCs for the $FABER_{height}$ and $FABER_{ROM}$ are slightly lower when compared to the study from Bagwell et al. (2018). The latter reported ICCs between 0.76 and 0.86 for intra-rater reliability and 0.68 for inter-rater reliability. The ICCs reported in our study are between 0.61 and 0.71 for $FABER_{height}$ and between 0.46 and 0.62 for $FABER_{ROM}$. The SSP condition helps to obtain the highest ICC in both variables (height and ROM). Inter-rater reliability were ranged from 0.46 to 0.58 for the $FABER_{height}$ and from 0.33 to 0.54 for $FABER_{ROM}$. Raters were asked to put a downward pressure on the tested leg until the opposite superior iliac spine moved under their opposite hand.

This is the first study to assess both, intra- and inter-rater reliability of $FADIR_{ROM}$. Although no comparison can be made with previous study, the intra-rater ICC obtained were ranged between 0.71 and 0.76 across the four conditions. No statistical different were observed between conditions. The inter-rater ICC were ranged between 0.40 and 0.57.

As expected, the intra-rater ICC values were higher than the inter-rater ICC values for both tests. Techniques used between raters may present some differences either to align the goniometer with the anatomical landmarks or to hold the dynamometer. This could also partly explain the observed variability between our study and previous studies. These results seem to highlight the importance of standardized procedures to limit variability.

Concerning the internal rotation ROM, ICCs are slightly lower when compared to previous studies. Krause et al. (2015) and Charlton et al. (2015) reported ICCs of 0.84 in a seated position. Charlton et al. (2015) also reported an ICC of 0.93 in a lying position. The seated position may help to stabilize the lower limb on the examination table. In lying position, the capsular ligaments are tight (hip extension) limiting multi axial movements. On the other hand, hip flexed at 90° decreases hip ligaments tension, that facilitates possible multi axial movement, which increasing variability in the internal rotation.

Standard error of measurement and minimal detectable difference

Concerning the influence of the controlled pressure duration for the intra-rater reliability, it helps to decrease the SEM and MDD for hip IR_{ROM} and $FADIR_{ROM}$ (C vs CPD). Concerning hip IR ROM, SEM and MDD decrease from 5.7° to 5.1° and from 15.9° to 14.1°, respectively. Concerning $FADIR_{ROM}$, SEM and MDD decrease from 5.8° to 5.6° and from 16.2° to 15.6°, respectively. SEM and MDD values for hip IR_{ROM} increase between SSP and M showing greater variability (SEM: 5.1° to 5.7°, MDD 14.2° to 15.9°). Unlike the hip IR_{ROM}, the dynamometer helps to decrease the variability between the SSP and M for $FADIR_{ROM}$.

The SEM and MDD values for $FABER_{height}$ increase from 1.4 cm to 1.8 cm and from 4.0 cm to 5.0 cm (C vs CPD). The SEM and MDD showed the same pattern between the SSP and M (SEM: 1.6 cm to 2.0 cm and MDD: 4.4 cm to 5.7 cm). According to our results, the SEM and MDD for $FABER_{ROM}$, increase from 3.4° to 3.7° and from 9.5° to 10.3°, respectively (C vs CPD). The SEM and MDD increase between the SSP and M (SEM: 2.9° to 3.5° and MDD 8.0 to 9.7). The classic condition should be prioritized for $FABER_{height}$ while the subject specific position should be used for $FABER_{ROM}$ assessment.

Our results showed an increase of SEM and MDD values for the $FABER_{ROM}$, $FABER_{height}$ and hip IR_{ROM} between the SSP and M (higher variability). Hypothetically, while the use of the dynamometer seems to decrease variability between the classic and controlled pressure duration conditions, the combination of the dynamometer and the subject specific position would increase the variability.

The SSP condition might brings problems in maintaining a steady rotational axis with var ying hip degrees of flexion or abduction. Thereby, modifying the hip position may modify the ligament properties (tightening or loosening) affecting the total range of motion (19, 20).

According to our results, the highest inter-rater reliability coefficients were present in the classic condition for $FABER_{height}$, $FABER_{ROM}$ and hip IR_{ROM}. The use of the dynamometer has improved the inter-rater ICC between the C and CPD only for the $FADIR_{ROM}$ (0.40 to 0.57).

The subject-specific position (SSP) showed one difference in the observed mean for the $FABER_{ROM}$. Smaller ROMs were observed in the conditions containing the subject-specific position (SSP and M) when compare to the conditions not involving this specificity (C and CPD). As stated previously, the osseous structure will adapt to the pattern of movements frequently used (12). Movements might had placed the hip in a position with greater osseous contact between the femoral head and acetabulum, decreasing the total ROM obtained.

Limitations

This study only used young asymptomatic participant. Participants with joint disease or pain may have resulted in different outcomes. Although it could have showed differences between asymptomatic and symptomatic participants, this would have also brought new limitations. Each participant was tested up to 24 times in two sessions. We though that the manifestation of pain in a symptomatic population might have brought variability and irritability among the participants.

Therefore, it is difficult to define the presence and type of FAI, thus bringing inter-subject variability (17,18). The use of the dynamometer might have brought some difficulties in the assessment of hip IR_{ROM} and $FADIR_{ROM}$. The dynamometer, by its size, made it difficult to hold the lower limb when performing the tests. This difficulty could reduce the accuracy of the test performed under the two conditions requiring its use. The assessment of the ROM needs acute sensory capacities to feel the end range of motion. A part of the clinician's attention might be to control the dynamometer thus decreasing the proprioceptive capacities. Lastly, the size of the leg may have brought some challenges to hold both, the leg and dynamometer.

Conclusion

We found that the CPD condition showed the highest ICCs for hip IR_{ROM} and $FADIR_{ROM}$, and that the SSP condition showed the highest ICCs for $FABER_{height}$ and $FABER_{ROM}$. Contrary to our hypothesis, the subject-specific position did not significantly modify the ROMs obtained. This might be explained by the fact that the participants were asymptomatic and presented no osseous modifications. The controlled pressure duration seems to improve reliability in IR_{ROM} and $FADDIR_{ROM}$, and future studies might evaluate this effect in a symptomatic population. $FABER_{ROM}$ and $FABER_{height}$ show, as in previous studies, low to moderate levels of intra-rater and interrater reliability, and this problem needs to be addressed for this provocative test. Lastly, levels of interrater reliability show that these tests should be performed by the same assessor to properly evaluate any change in the individual's condition between different assessment periods.

References

1. Van Trijffel, E., van de Pol, R. J., Oostendorp, R. A., & Lucas, C. Inter-rater reliability for measurement of passive physiological movements in lower extremity joints is generally low: a systematic review. J Physiother, 2010; 56(4), 223-235.

2. Tijssen, M., van Cingel, R., Willemsen, L., & de Visser, E. Diagnostics of femoroacetabular impingement and labral pathology of the hip: a systematic review of the accuracy and validity of physical tests. Arthroscopy, 2012; 28(6), 860-871.

3. Martin, H. D., Kelly, B. T., Leunig, M., Philippon, M. J., Clohisy, J. C., Martin, R. L.et al. The pattern and technique in the clinical evaluation of the adult hip: the common physical examination tests of hip specialists. Arthroscopy, 2010; 26(2), 161-172.

4. Altman, R., Alarcon, G., Appelrouth, D., Bloch, D., Borenstein, D., Brandt, K. et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Athritis Rheum, 1991; 34(5), 505-514.

5. Gabbe, B. J., Bennell, K. L., Wajswelner, H., & Finch, C. F. Reliability of common lower extremity musculoskeletal screening tests. Phys Ther Sport, 2004; 5(2), 90-97.

6. Kouyoumdjian, P., Coulomb, R., Sanchez, T., & Asencio, G. Clinical evaluation of hip joint rotation range of motion in adults. Orthop Traumatol Surg Res, 2012; 98(1), 17-23.

7. Krause, D. A., Hollman, J. H., Krych, A. J., Kalisvaart, M. M., & Levy, B. A. Reliability of hip internal rotation range of motion measurement using a digital inclinometer. Knee Surg Sport Tr A, 2015; 23(9), 2562-2567.

8. Philippon, M. J., Ho, C. P., Briggs, K. K., Stull, J., & LaPrade, R. F. "Prevalence of increased alpha angles as a measure of cam-type femoroacetabular impingement in youth ice hockey players." Am J Sports Med, 2013; 41(6): 1357-1362.

9. Bagwell, J. J., Bauer, L., Gradoz, M., & Grindstaff, T. L. "The Reliability of Faber Test Hip Range of Motion Measurements." Int J Sports Phys Ther, 2016; 11(7): 1101-1105.

10. Agricola, R., Heijboer, M. P., Ginai, A. Z., Roels, P., Zadpoor, A. A., Verhaar, J. A., & al. A cam deformity is gradually acquired during skeletal maturation in adolescent and young male soccer players: a prospective study with minimum 2-year follow-up. Am J Sports Med, 2014; 42(4), 798-806.

11. Lahner, M., Walter, P. A., von Schulze Pellengahr, C., Hagen, M., von Engelhardt, L. V., & Lukas, C. Comparative study of the femoroacetabular impingement (FAI) prevalence in male semiprofessional and amateur soccer players. Arch Orthop Trauma Surg, 2014; 134(8), 1135-1141.

12. Zadpoor, A. A. "Etiology of Femoroacetabular Impingement in Athletes: A Review of Recent Findings." Sports Med, 2015; 45(8): 1097-1106.

13. Ayeni, O. R., Banga, K., Bhandari, M., Maizlin, Z., Golev, D., Harish, S., & al. "Femoroacetabular impingement in elite ice hockey players." Knee Surg Sports Traumatol Arthrosc, 2014; 22(4): 920-925.

14. Furness, J., Schram, B., Cox, A. J., Anderson, S. L., & Keogh, J. Reliability and concurrent validity of the iPhone® Compass application to measure thoracic rotation range of motion (ROM) in healthy participants. Peer J, 2018, 6, e4431.

15. Philippon, M. J., Maxwell, R. B., Johnston, T. L., Schenker, M., & Briggs, K. K. "Clinical presentation of femoroacetabular impingement." Knee Surg Sports Traumatol Arthrosc, 2007; 15(8): 1041-1047.

16. Narvani, A. A., Tsiridis, E., Kendall, S., Chaudhuri, R., & Thomas, P. "A preliminary report on prevalence of acetabular labrum tears in sports patients with groin pain." Knee Surg Sports Traumatol Arthrosc, 2003; 11(6): 403-408.

17. Kapron, A. L., Aoki, S. K., Peters, C. L., & Anderson, A. E. "In-vivo hip arthrokinematics during supine clinical exams: Application to the study of femoroacetabular impingement." J Biomech, 2015; 48(11): 2879-2886.

18. Wyss, T. F., Clark, J. M., Weishaupt, D., & Nötzli, H. P. Correlation between internal rotation and bony anatomy in the hip. Clin Orthop Relat Res 2007; 460, 152-158.

19. Martin, H. D., Savage, A., Braly, B. A., Palmer, I. J., Beall, D. P., & Kelly, B. The function of the hip capsular ligaments: a quantitative report. Arthroscopy, 2008; 24(2), 188-195.

20. van Arkel, R. J., Amis, A. A., & Jeffers, J. R. The envelope of passive motion allowed by the capsular ligaments of the hip. J Biomech, 2015; 48(14), 3803-3809.

21. Epstein, D. M., McHugh, M., Yorio, M., & Neri, B.. Intra-articular hip injuries in National Hockey League players: a descriptive epidemiological study. Am J Sports Med, 2013, 41(2), 343-348.

22. Prather, H., Harris-Hayes, M., Hunt, D. M., Steger-May, K., Mathew, V., & Clohisy, J. C. Reliability and agreement of hip range of motion and provocative physical examination tests in asymptomatic volunteers. PM & R, 2010 2(10), 888-895.

23. Tijssen, M., van Cingel, R. E. H., de Visser, E., Hölmich, P., & Nijhuis-van der Sanden, M. W. G. Hip joint pathology: relationship between patient history, physical tests, and arthroscopy findings in clinical practice. Scand J Med Sci Sports, 2017, 27(3), 342-350.