Physical Therapy in Sport 38 (2019) 115–122

Contents lists available at ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Original Research

Knee and hip strength measurements obtained by a hand-held dynamometer stabilized by a belt and an examiner demonstrate parallel reliability but not agreement

Lidiane L. Florencio [a, *], Jaqueline Martins [b], Marcelo R.B. da Silva [b], Janaina R. da Silva [b], Gustavo L. Bellizzi [b], Débora Bevilaqua-Grossi [b]

[a] Physiotherapy, Occupational Therapy, Physical Medicine and Rehabilitation Department, Universidad Rey Juan Carlos, Madri, Spain
[b] Department of Health Sciences, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil

A R T I C L E I N F O

Article history:
Received 18 February 2019
Received in revised form
18 April 2019
Accepted 19 April 2019

Keywords:
Muscle strength
Knee
Hip
Reproducibility of results

A B S T R A C T

Objectives: To verify the intrasession reliability and the agreement between strength measurement of hip and knee muscles using hand-held dynamometer stabilized by a belt or by an examiner.
Design: Test-retest design.
Setting: Knee and hip muscles strength were measured bilaterally using hand-held dynamometer stabilized by a belt and by an examiner.
Participants: 24 young and healthy participants.
Main outcome measures: The reliability was verified by the intraclass correlation coefficient$_{2,1}$ and the standard error of measurement. Agreement between stabilization methods was verified by the Bland Altman's method.
Results: Reliability was excellent for all muscle groups when stabilized by a belt (intraclass correlation coefficient = 0.78 to 0.95) and by the examiner (intraclass correlation coefficient = 0.83 to 0.97); standard error of measurement ranged between 1kgf to 4kgf at both methods, but they are proportionally lower when stabilized by the examiner. No agreement between both methods was identified for all knee strength measurements and for bilateral hip flexion, right internal and external rotation and left adduction.
Conclusions: The hand-held dynamometer is reliable for hip and knee strength evaluation despite of the stabilization method. However, for the majority of the movements, greater strength and lower error are expected when the examiner stabilizes it.

1. Introduction

Muscle strength measurement is an important clinical aspect when proposing a treatment, to measure a therapy's effectiveness or when considering patient discharge criteria. In sports like track and field, football and ice hockey the assessment of both hip and knee strength seems to be relevant to quantify muscle strength deficits and predict injury (Thorborg, Bandholm, & Hölmich, 2013).

Muscle strength has been frequently evaluated in scientific contexts and clinical practice using a hand-held dynamometer (HHD), stabilized by a belt or examiner, which analyzes strength in isometric muscle contraction and is a low-cost and easily portable device (Halabchi, Mazaheri, & Seif-Barghi, 2013; Krause et al., 2014). However, as for any evaluation tool, whether stabilized by a belt or the examiner, it has to be reliable and validated. For instance, it still unclear which method of stabilization would be more appropriated.

Studies show that a HHD stabilized by the examiner is a reliable tool for lower limb muscle strength evaluation (0.70–0.98) (Chamorro, Armijo-Olivo, De La Fuente, Fuentes, & Javier Chirosa, 2017; Fulcher, Hanna, & Raina Elley, 2010; Ieiri et al., 2015; Kelln, McKeon, Gontkof, & Hertel, 2008; Kim & Lee, 2015; Kim, Kim,

* Corresponding author. Department of Physiotherapy, Occupational Therapy, Rehabilitation, and Physical Medicine, Universidad Rey Juan Carlos, Campus de Alcorcón, Avenida Atenas s/n, Alcorcón, Madrid, 28922, Spain.

E-mail addresses: lidianelimaflorencio@gmail.com (L.L. Florencio), jaqueline_mh@yahoo.com.br (J. Martins), marcelo09_93@hotmail.com (M.R.B. da Silva), janainaknak@hotmail.com (J.R. da Silva), glbellizzi.fisio@hotmail.com (G.L. Bellizzi), deborabg@fmrp.usp.br (D. Bevilaqua-Grossi).

https://doi.org/10.1016/j.ptsp.2019.04.011
1466-853X/

Seo, & Kang, 2014; Krause et al., 2014; Mentiplay et al., 2015; Scott, Bond, Sisto, & Nadler, 2004). However, criticism revolves around the potential influence of the examiner in the measurements and the difficult to stabilize the HHD in stronger subjects (Bohannon, Kindig, Sabo, Duni, & Cram, 2012; Ieiri et al., 2015; Kelln et al., 2008; Scott et al., 2004; Wadsworth, Nielsen, Corcoran, Phillips, & Sannes, 1992; Wikholm & Bohannon, 1991). As consequence, belts were suggested for dynamometer stabilization (Ieiri et al., 2015), expecting that its use in HHD testing could increase the device's reliability for knee and hip strength.

Indeed, the reliability of belt-stabilized HHD for knee and hip muscle groups have been demonstrated to be moderate to excellent (0.49–0.99) in the evaluation of subjects with no injury or athletes (Bohannon, Chu, & Portz, 2015; Bohannon, Pritchard, & Glenney, 2013; Hansen, McCartney, Sweeney, Palimenio, & Grindstaff, 2015; Ieiri et al., 2015; Kim et al., 2014; Martins, da Silva, da Silva, & Bevilaqua-Grossi, 2017; Oliveira, Pilz, Santos; Junior, Vasconcelos, Mello, & Grossi, 2018; Thorborg et al., 2013; Toonstra & Mattacola, 2013). Moreover, excellent reliability (0.93–0.98) has also been reported when PVC pipe device stabilizes the HHD (Jackson, Cheng, Smith Jr, & Kolber, 2017).

Although HHD reliability has already been analyzed using different stabilization methods, they differ on parameters of the examiner's position and the participant's position during the test. Therefore, it is still not clear whether the belt-stabilized method is more reliable than the method stabilized by the examiner in the lower limb evaluation using the same test protocol. It is also not clear if these measures agree with each other or if there is a systematic error.

Accordingly, this study aimed to verify the agreement of lower limb strength measurements using HHD stabilized by a belt or by the examiner in young and health participants. It also aimed to describe the intrasession reliability of both methods for knee flexion and extension movements, and adduction, abduction, flexion, extension, and internal and external rotation of the hip. We hypothesized that a HHD stabilized by a belt would be more reliable and that the examiner stabilizing a HHD would have an influence on the test, generating higher error values for the measurement.

2. Methods

2.1. Participants

Twenty-four participants (12 male and 12 female) without reported hip or knee dysfunction, aged 18–28 years (mean age: 23.1 ± 3 years, mean height: 170 ± 10 cm, and mean weight: 68.5 ± 13 kg), participated in this study. The participants were recruited from an academic setting and signed an informed consent form approved by the local Ethics Committee (Case No. 15917/2014).

Women included in the study were out of their menstrual period, or one week before, due to alterations in the motor response in this period (Barbosa, Montebelo, & Guirro, 2007). Participants reporting pain, those with history of injury or orthopedic surgery in the lower limbs in the previous six months, and those with neurological disorders were excluded from the study (Kollock, Onate, & Van Lunen, 2010). Participants considered athletes were also excluded based on the International Physical Activity Questionnaire (IPAQ) – short version (Matsudo et al., 2002).

The sample size calculation was conducted for reliability data considering 2 observations per volunteer ($k = 2$), an intraclass correlation coefficient (ICC) estimated at 0.80, an amplitude-based confidence interval of 0.3, that is, $0.5 \leq ICC \leq 1.0$ and a confidence coefficient with an 0.05 alpha level, resulting in 24 volunteers (Bonett, 2002).

2.2. Examiner

The examiner was a 22-year-old male, undergraduate at Physical Therapy, with a body mass index of 22.78 kg/m^2. He took a 30-h training to become familiar with the Lafayette® HHD (model 01163, Lafayette® Instrument Company, IN, USA) in tests stabilized by a belt or examiner.

The isometric strength evaluation of examiner was conducted in the examiner dominant limb to measure the hand grip strength using a Jamar® dynamometer, model BL5001 (Jamar® Hydraulic Hand Dynamometer, Santa Ana, CA 92705), indicating a hand grip strength of 50Nm. Upper limb strength was measured using an isokinetic dynamometer (Biodex 3®, Biodex Medical Systems Inc., Shirley, NY, USA). The mean strength observed for extensors and flexors of the elbow were: 58.7Nm and 85.1Nm; and for extensors and flexors of the shoulder were: 94.3Nm and 58.4 Nm.

2.3. Instruments

The tests were conducted with a Lafayette® HHD stabilized by Velcro® belts or by the examiner. Additional belts were used to stabilize the participant and minimize potential compensations of the trunk, hip and knee.

The belt that stabilized the Lafayette HHD had its female face glued to the posterior portion of the HHD, preventing belt displacement on the dynamometer during the tests. A 3 mm thermoplastic board covered by a 2 mm EVA plastic was used to adapt better to the thigh and a shinguard to adjust better to the leg (Figs. 1 and 2). Hansen et al. [24] concluded that a surface between the participant's leg and the HHD increases the comfort and produces strength results of knee extensors closer to values found by the isokinetic dynamometer. Fig. 1 shows the test of hip flexors with a HHD stabilized by the examiner and by a belt.

2.4. Procedures

The eligible participants were evaluated for anthropometric information of body weight (kg) and height (m) and provided personal information. Bilateral muscle strength tests for the flexor and extensor muscle groups of the knee, and the flexor, extensor, adductor, abductor, and internal and external rotator muscle groups of the hip were conducted with a Lafayette® HHD stabilized by a belt and by an examiner.

To optimize the assessment time and minimize the participant's muscle fatigue, two sets of tests were proposed: one started from the knee joint (knee extension; hip flexion; knee flexion; and internal rotation, external rotation, extension, abduction, and adduction of the hip) and one started from the hip joint (abduction, adduction, extension, external rotation, and internal rotation of the hip; knee flexion; hip flexion; and knee extension). Two draws were conducted, the first to determine the order of the joint to be

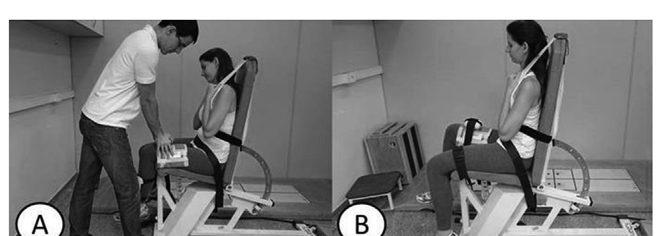

Fig. 1. Tests of hip flexors with hand-held dynamometer stabilized by the examiner (A) and by a belt (B). (1 column).

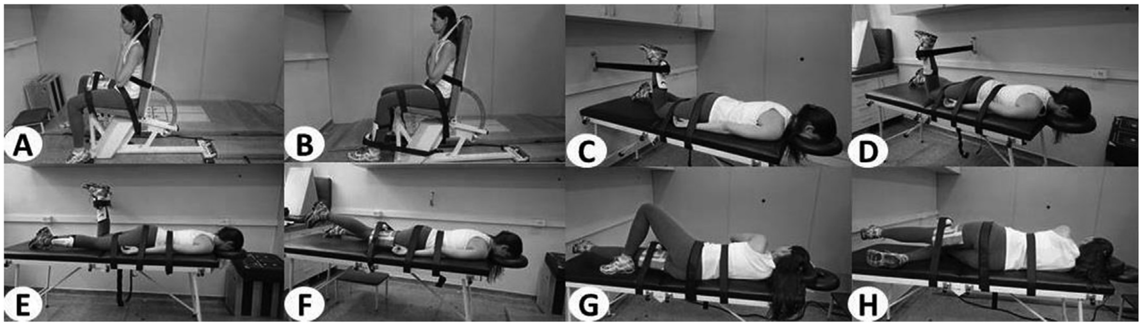

Fig. 2. Hand-held dynamometer stabilized by a belt for hip flexion (A), knee extension (B), knee flexion(C) and for hip internal rotation (D), external rotation (E) extension (F), adduction (G) and abduction (H). (2 columns).

test and the second to determine if the tests would be conducted first with the right or left lower limb. The draw was conducted using sealed thick envelopes randomly arranged on a table. After the draw, participants were submitted to a 10-min warm-up period using an exercise bike with no load, and three 5-s isometric contractions were performed for each muscle group, with 30-s intervals to rest between repetitions and 5-min intervals for each muscle group analyzed (Toonstra & Mattacola, 2013). All individuals were encouraged to perform the tests at maximum voluntary effort and received a verbal encouragement. Assessment with belt-stabilization method was performed on the first day and the examiner stabilization method was repeated within 3–5 days. Tests were schedule in the same period of the day and the order of tests was the same for both stabilization methods.

The tests were conducted for knee extensors and hip flexors with individuals sitting on a leg extension machine. Knee flexors and hip extensors and rotators were evaluated in a prone position, and the hip adductors and abductors in a side lying position on a stretcher (Fig. 2).

2.5. Statistical analysis

The statistical analyses were processed by SPSS software, version 17 (SPSS Inc., Chicago, IL, USA) adopting a significance level of 0.05. Intrasession reliability of strength values generated by the HHD stabilized by a belt or examiner was calculated using the $ICC_{2,1}$ associated with a 95% confidence interval. The reliability was classified as poor (<0.40), moderate (<0.40 and < 0.75) or excellent (>0.75) (Fleiss, 1986). The measurement error was analyzed through the standard error of measurement, with standard deviation (SD) obtained from 3 repetitions by the formula “Standard error of measurement $= SD\sqrt{1\text{-}ICC}$”. For clinical purposes, the proportional standard error of measurement was calculated as percentage of the mean measured with each stabilization method.

The agreement between HHD stabilized by a belt and by the examiner was verified by the Bland-Altman method. A graph was constructed fo each direction, displaying the mean difference (examiner minus belt stabilization) and the 95% limits of agreement (LOA) (Bland & Altman, 1999). Mean difference, which represents the bias, was calculated, used for the limits of agreement estimation and test by one sample *t*-test. Significant bias is suggested when the 95% confidence interval of the mean difference does not include the line of equality (bias = 0) (Giavarina, 2015). Limits of agreement was calculated by the formula “LOA = bias ±1.96*SD” (Bland & Altman, 1999). Linear regressions between the difference and the mean measure for each movement assessed were conducted to analyze proportional bias (Bland & Altman, 1999).

3. Results

HHD stabilized by a belt demonstrated excellent intrasession reliability for bilateral strength tests (ICC = 0.78 to 0.95) for the hip and knee muscle groups (Table 1). The assessment of knee muscle group presented standard error of measurement values from 1.2kgf to 3.8kgf (8%–11%). In the hip muscle group, the standard error of measurement was 1.0kgf to 3.6kgf (7%–15%).

Intrasession reliability of the HHD stabilized by the examiner was also excellent (ICC = 0.83 to 0.97) (Table 2). For the knee muscle group, the standard error of measurement was 1.0kgf to 4.1kgf (7%–9%). While for the hip muscle group, the standard error of measurement was 0.7kgf to 2.3kgf (6%–9%).

Comparing the standard error of measurement observed for each joint, in knee muscle groups, it was higher for extensors, but the proportional standard error of measurement was similar to that observed for flexors, for both methods. In hip muscle groups, for both methods, internal and external rotators presented the lowest standard error of measurement in relation to other muscles.

Table 1
Mean strength (kgf), intrassesion reliability and standard error of measurement of hand-held dynamometer stabilized by a belt (n = 24).

Muscle Group	Right			Left		
	Mean (SD)	$ICC_{2,1}$ (CI 95%)	SEM Kgf/%SEM	Mean (SD)	$ICC_{2,1}$ (CI 95%)	SEM Kgf/%SEM
Knee extensors	41.5 (14.2)	0.93 (0.86; 0.97)	3.8/9%	41.6 (15.5)	0.95 (0.90; 0.98)	3.5/8%
Knee flexors	12.7 (4.6)	0.93 (0.88; 0.97)	1.2/10%	12.5 (4.1)	0.90 (0.82; 0.95)	1.3/11%
Hip internal rotators	10.2 (3.4)	0.91 (0.84; 0.96)	1.1/10%	10.1 (3.3)	0.90 (0.81; 0.95)	1.0/10%
Hip external rotators	11.1 (3.8)	0.93 (0.87; 0.97)	1.0/9%	11.1 (4.0)	0.90 (0.82; 0.95)	1.3/12%
Hip extensors	28.8 (10.3)	0.94 (0.89; 0.97)	2.5/9%	26.9 (9.8)	0.93 (0.86; 0.96)	2.6/10%
Hip flexors	24.7 (9.1)	0.93 (0.88; 0.97)	2.4/10%	23.6 (8.5)	0.92 (0.86; 0.96)	2.4/10%
Hip adductors	21.3 (7.0)	0.86 (0.75; 0.93)	2.7/13%	20.0 (5.8)	0.78 (0.62; 0.89)	2.9/15%
Hip abductors	30.4 (9.6)	0.95 (0.90; 0.97)	2.1/7%	30.4 (9.6)	0.87 (0.76; 0.93)	3.6/12%

SD: standard deviation; ICC, intraclass correlation coefficient; SEM, standard error of measurement; CI, confidence interval; % SEM, a percentage of the mean strength generated by the muscle group.

Table 2
Mean strength (kgf), intrassesion reliability and standard error of measurement of hand-held dynamometer stabilized by the examiner (n = 24).

Muscle Group	Right			Left		
	Mean (SD)	$ICC_{2,1}$ (CI 95%)	SEM Kgf/%SEM	Mean (SD)	$ICC_{2,1}$ (CI 95%)	SEM Kgf/%SEM
Knee extensors	46.5 (14.0)	0.93 (0.87; 0.97)	3.7/8%	46.3 (15.2)	0.93 (0.87; 0.97)	4.1/9%
Knee flexors	14.3 (6.1)	0.97 (0.94; 0.99)	1.1/7%	14.6 (5.7)	0.97 (0.94; 0.99)	1.0/7%
Hip internal rotators	11.4 (3.2)	0.92 (0.86; 0.96)	.9/8%	10.7 (2.4)	0.90 (0.83; 0.96)	0.8/7%
Hip external rotators	12.0 (3.4)	0.95 (0.89; 0.97)	.8/6%	11.5 (3.5)	0.96 (0.90; 0.98)	0.7/6%
Hip extensors	27.4 (6.5)	0.88 (0.78; 0.94)	2.3/9%	27.4 (6.5)	0.90 (0.82; 0.95)	2.1/8%
Hip flexors	30.6 (9.5)	0.96 (0.91; 0.98)	1.9/6%	30.0 (9.7)	0.96 (0.93; 0.98)	1.9/6%
Hip adductors	22.4 (6.0)	0.92 (0.85; 0.96)	1.7/8%	23.0 (6.9)	0.92 (0.84; 0.96)	2.0/9%
Hip abductors	29.1 (5.3)	0.83 (0.69; 0.91)	2.3/8%	30.2 (6.1)	0.88 (0.78; 0.94)	2.2/7%

SD: standard deviation; ICC, intraclass correlation coefficient; SEM, standard error of measurement; CI, confidence interval; % SEM, a percentage of the mean strength generated by the muscle group.rowhead

However, the proportional standard error of measurement differed with the method. For the belt-stabilized method, the proportional standard error of measurement was lower for hip extensors and higher for hip abductors; while for the method stabilized by the examiner, the proportional standard error of measurement was lower for external rotators and flexors of the hip and similar to that observed for other muscles.

Significant bias was identified between both methods at all knee strength measures (Fig. 3), since the line of equality (bias = 0) is not within the confidence interval of the mean difference. Measurements registered with the examiner stabilization were generally greater than those observed for belt-stabilization. Mean differences observed were: 5.0kgf (95%CI: 1.6 to 8.4; $P = .006$) for right extension; 4.7kgf (95%CI: 1.1 to 8.3; $P = .012$) for left extension; 1.6kgf (95% CI:0.1 to 3.3; $P = .049$) for right flexion and 2.1kgf (95%CI: 0.6 to 3.6; $P = .008$) for left flexion. Additionally, proportional biases were identified at right ($\beta = 0.31$; $P = .048$) and left knee flexion ($\beta = 0.36$; $P = .020$), suggesting that for stronger the participants greater errors could be expected.

For hip strength measurements, significant bias was identified

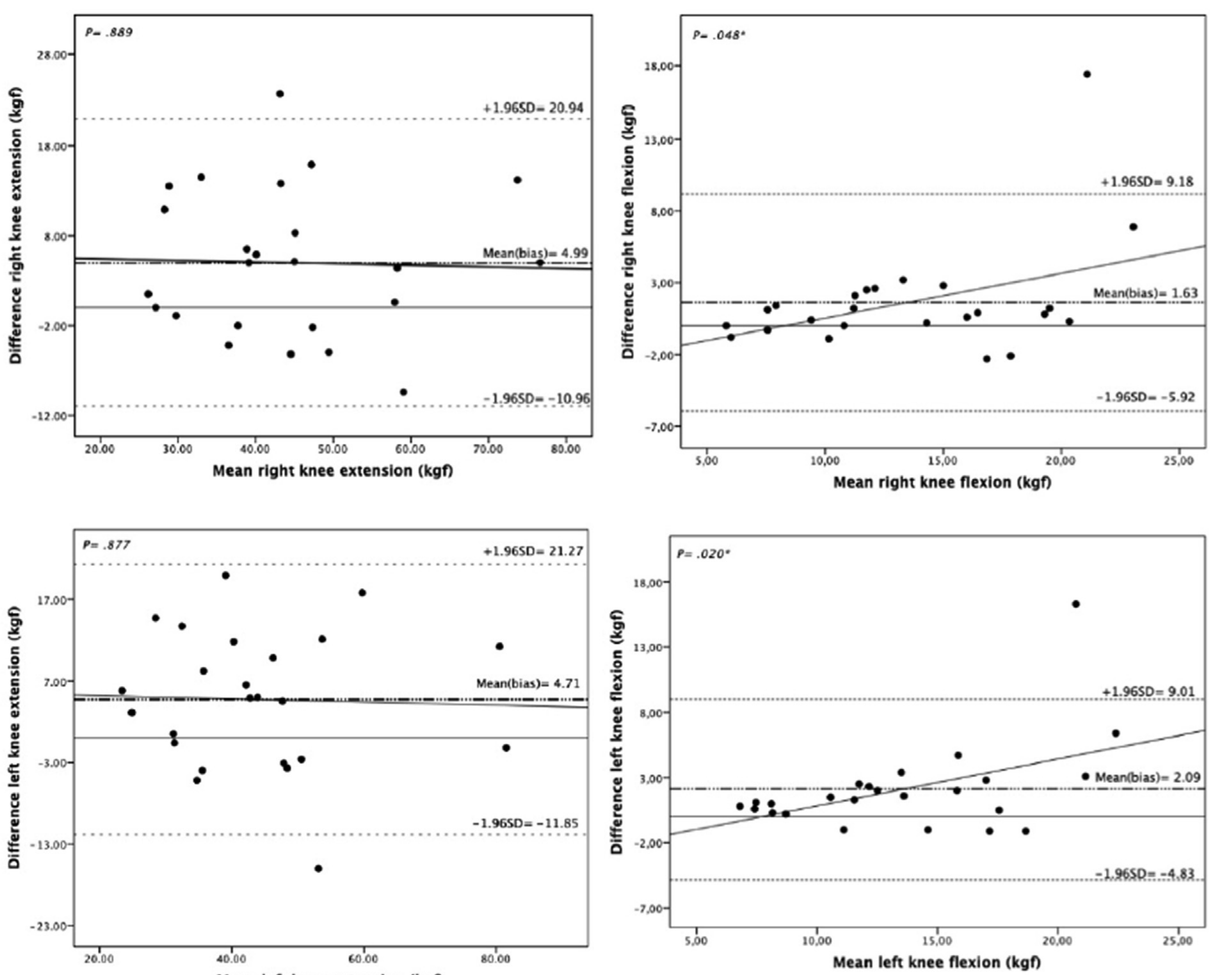

Fig. 3. Bland-Altman plots comparing the belt-stabilized and the examiner-stabilized hand-held dynamometer in assessing strength of knee extensors (A and B) and flexors (C and D). The P value is for the slope of the regression line. Abbreviation: SD, standard deviation. (1 column).

for right internal rotation and right external rotation (Fig. 4), bilateral flexion (Fig. 5) and left adduction (Fig. 6). Similar to the pattern observed for knee, at these movements with significant biases, measurements stabilized by the examiner generally greater values than those stabilized by the belt. Mean differences observed were: 1.3kgf (95%CI: 0.6 to 1.9; $P = .001$) for right internal rotation and 0.9kgf (95%CI: 0.1 to 1.7; $P = .027$) for right external rotation; 5.9kgf (95%CI: 2.8 to 9.0; $P = .001$) for right hip flexion; 6.4kgf (95% CI: 3.0 to 9.8; $P = .001$) for left hip flexion and 3.0kgf (95%CI: 1.2 to 4.9; $p = .003$) for left adduction. Proportional biases were identified at bilateral hip extension (right $\beta = -0.49$; left $\beta = -0.45$), bilateral abduction (right $\beta = -0.66$; left $\beta = -0.54$) and left internal rotation ($\beta = -0.33$). However, none of the significant proportional biases matched with those hip strength measures that present significant mean difference between both methods and as can be observed at Figs. 4–6 the error is expected to be greater at the upper and lower values of the strength measures.

4. Discussion

Strength tests conducted with the HHD stabilized by a belt or by an examiner are reliable for all muscle groups of the knee and hip. In addition, the belt-stabilized method presented measurement errors slightly higher than those for the method stabilized by the examiner. Despite of their excellent reliability, no agreement was identified between measures obtained by both stabilization protocols for all measurements of knee strength and the majority of hip strength.

Current reliability findings contradict our initial hypothesis that greater reliability and lower errors would be achieve for the belt-stabilization methods. It was hypothesized that using the belt rather than examiner stabilization would provide better stability to the HHD and it also would avoid the addition of the examiner force (Bohannon et al., 2012; Ieiri et al., 2015; Kelln et al., 2008; Scott et al., 2004; Wadsworth et al., 1992; Wikholm & Bohannon, 1991). However, in terms of reliability, it seems that these potential benefits do not impact enough in the consistency of the measurement to minimize substantially the measurement error. This is in accordance with Krause et al. (2014) that demonstrated that the reliability of strength measured by a HHD stabilized by the examiner was not influenced by the examiner's strength when testing hip extension, abduction and external rotation of healthy participants.

The excellent reliability described here for tests conducted with HHD stabilized by an examiner agree with the literature, which has also reported excellent reliability for the method stabilized by the examiner for the muscle groups of extensors (ICC = .87 to .94) (Kim & Lee, 2015; Krause et al., 2014; Mentiplay et al., 2015), flexors (ICC = 0.81 to 0.97) (Kim & Lee, 2015; Krause et al., 2014; Mentiplay et al., 2015; Scott et al., 2004), abductors (ICC = 0.82 to 0.94) (Kelln et al., 2008; Kim & Lee, 2015; Krause et al., 2014), and external rotators of the hip (ICC = 0.85 to 0.97) (Kelln et al., 2008; Kim & Lee,

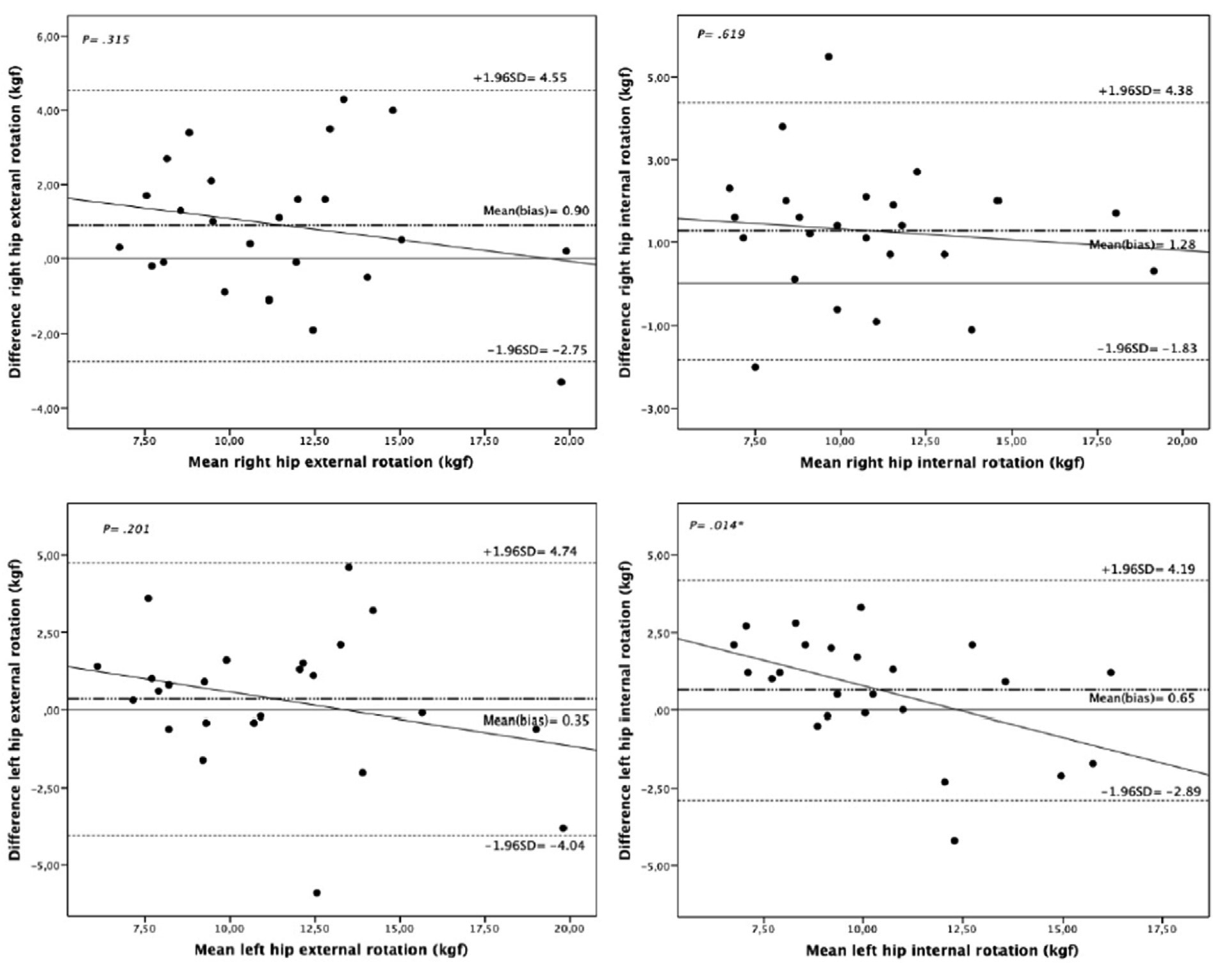

Fig. 4. Bland-Altman plots comparing the belt-stabilized and the examiner-stabilized hand-held dynamometer in assessing strength of hip muscles to perform external (A and B) and internal rotation (C and D). The P value is for the slope of the regression line. Abbreviation: SD, standard deviation. (1 column).

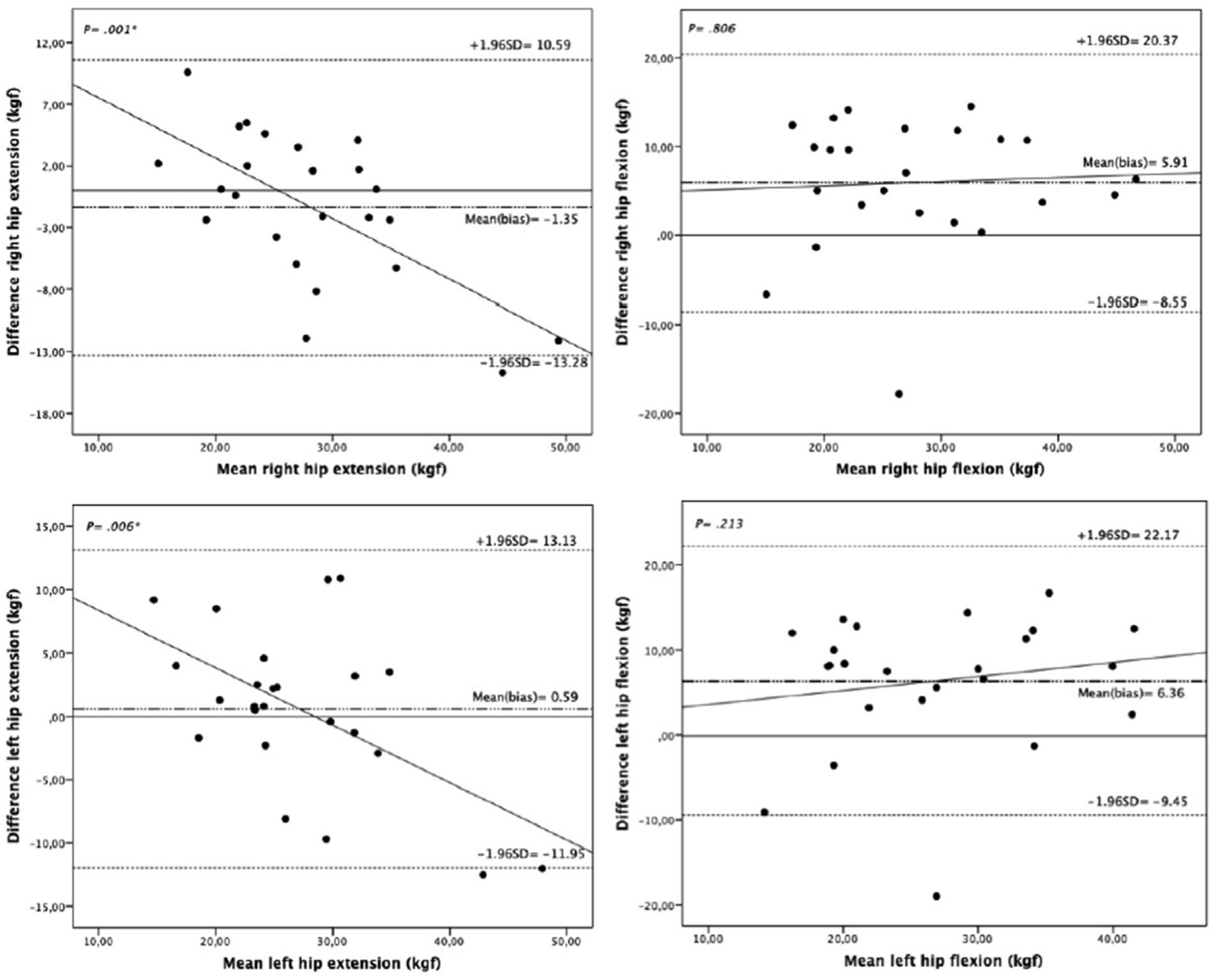

Fig. 5. Bland-Altman plots comparing the belt-stabilized and the examiner-stabilized hand-held dynamometer in assessing strength of hip muscles to perform extension (A and B) and flexion (C and D). The P value is for the slope of the regression line. Abbreviation: SD, standard deviation. (1 column).

2015; Krause et al., 2014), and for knee extensors (ICC = 0.98) (Kim & Lee, 2015; Krause et al., 2014) and flexors (ICC = 0.83 to 0.98) (Kelln et al., 2008; Kim & Lee, 2015; Krause et al., 2014; Mentiplay et al., 2015) of healthy participants.

However, for the belt-stabilized method, the excellent reliability agree partially with literature as excellent reliability have been reported for knee extensors (ICC = 0.76 to 0.99) (Hansen et al., 2015; Kim et al., 2014; Martins et al., 2017; Toonstra & Mattacola, 2013), hip abductors (ICC = 0.80 to 0.99) (Ieiri et al., 2015; Martins et al., 2017; Oliveira et al., 2018), adductors (0.88–0.94) (Martins et al., 2017), flexors (0.76–0.96) (Martins et al., 2017; Oliveira et al., 2018), extensors (0.86–0.95) (Martins et al., 2017; Oliveira et al., 2018), and external rotators (0.80–0.90) (Martins et al., 2017). On the other hand, moderate reliability have been reported for knee flexors (ICC = 0.49 to 0.66) (Martins et al., 2017; Toonstra & Mattacola, 2013) and moderate-to-excellent reliability for internal rotators (0.70–0.80) (Martins et al., 2017).

Thus, the HHD is a reliable option in clinical evaluation of lower limb strength independently of how it is stabilized. However, the standard error of measurement and proportional standard error of measurement values should also be taken into account, as high error values may lead the examiner to observe increased muscle strength that is not true. Our study showed, for both methods for knee extensors and flexors, a dynamometer error of around 4kgf and 1kgf, respectively; for the hip, the error was approximately 1kgf for the rotators and 2kgf to 4kgf for the other muscle groups.

Surprisingly, the standard error of measurement and proportional standard error of measurement values were slightly lower in general for the method stabilized by the examiner. Among the few studies that analyzed the error measurements by comparing the methods stabilized by a belt and by an examiner (Ieiri et al., 2015; Kim et al., 2014), only absolute standard error of measurement have been reported between examiners at both methods. Similar standard error of measurement have been reported to the knee extensors in the sitting position (2.1 Nm to 2.6 Nm) (Kim et al., 2014) and, for hip abductors, lower standard error of measurement were described for the method stabilized by the examiner (6.7N–6.9N) compared to the method stabilized with a belt (7.7N–8.8N) for 2 of the 3 examiners of the study (Ieiri et al., 2015). It is possible that the greater error values for the belt-stabilized method, particularly the proportional standard error of measurement values, suggest insufficient dynamometer stabilization with belts only, that is, a better control of the dynamometer position is guaranteed when applied by the examiner.

On the other hand, it should be recognized that greater strength measures are generally expected when the examiner stabilizes the HHD, especially for strength measurement of knee muscles and for the hip flexion, abduction, internal and external rotation. Additionally, according to significantly proportional biases, greater differences between both methods are expected for stronger

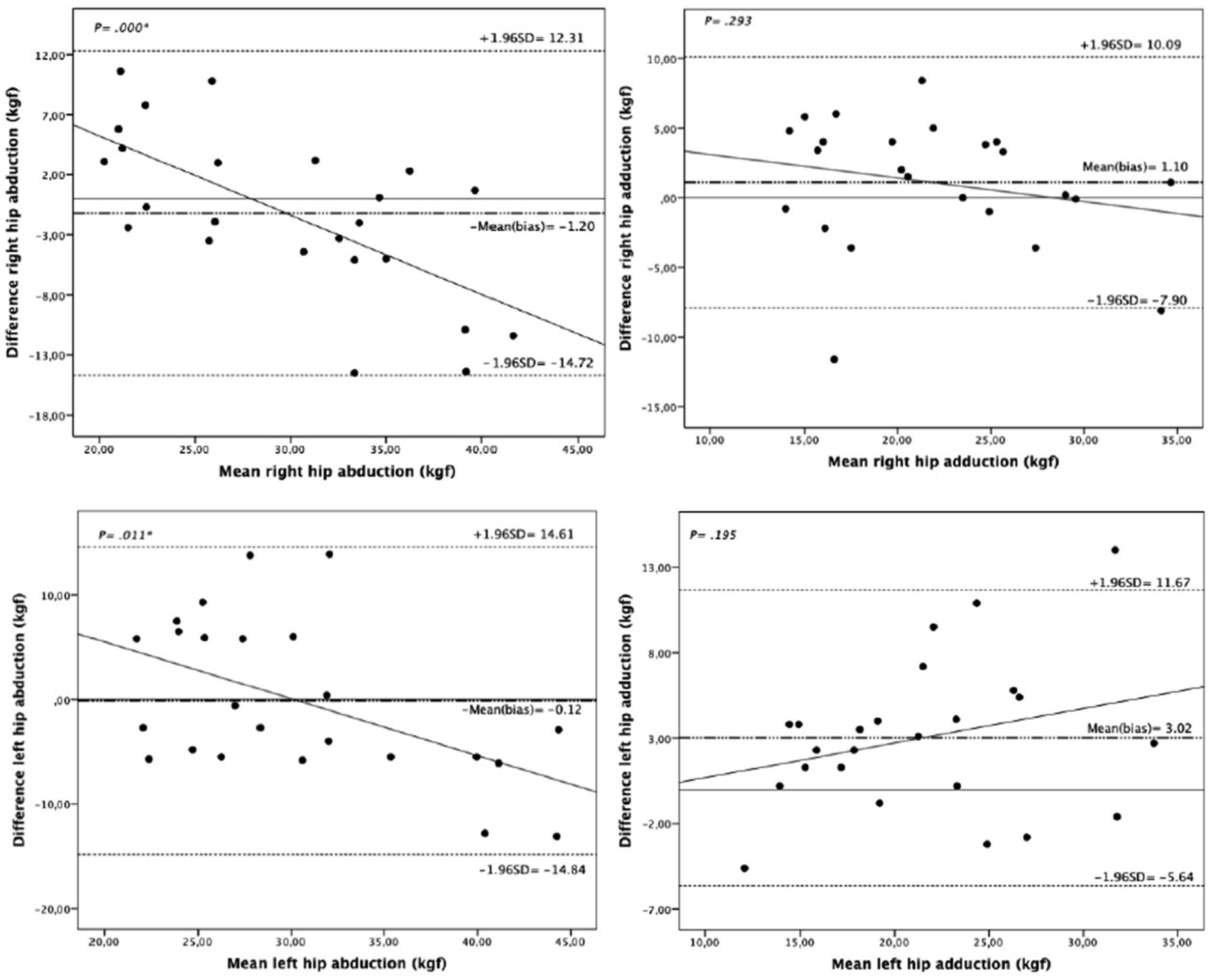

Fig. 6. Bland-Altman plots comparing the belt-stabilized and the examiner-stabilized hand-held dynamometer in assessing strength of hip muscles to perform abduction (A and B) and adduction (C and D). The P value is for the slope of the regression line. Abbreviation: SD, standard deviation. (1 column).

participants at knee flexion strength measurement.

The LOA between two stabilization methods reinforce that we cannot assume an agreement between them. Although there is no standard cut off or an acceptable limit to determine whether they agree or not (Giavarina, 2015), we can critically draw attention to LOA of all movements assessed. In the current data, upper LOA presented a greater magnitude than the lower limit in all cases. Considering the raw data, upper LOA ranged from about 4kgf (bilateral hip internal rotation, Fig. 4C and D) to 22kgf (left hip flexion, Fig. 5D). For a clinical perspective, it is worth to note that, in the best scenario, the upper LOA could be proportional to 35% of the mean measured in the participants, which is a 4kgf of upper LOA to a mean of 11.4kgf measured with the examiner stabilization methods. Moreover, in the worse scenario, upper LOA could be proportional to 70% of the mean, which is observed for knee flexion with an upper LOA of about 9kgf and a mean of 12.7kgf measured by the belt-stabilization method. This proportion would not be an acceptable margin in daily practice.

Therefore, the current findings support that both stabilization methods are indicated to clinical practice since both methods present excellent reliability. However, lower error values are expected when the examiner stabilizes the HHD, with proportional standard error of measurement lower than the 10% cut-off reference, which is generally considered to be a clinically relevant improvement or deterioration at clinical practice (Chamorro et al., 2017; Prentice & Kaminski, 2004). Moreover, measurements performed with both methods are not reasonably equivalent, so care should be taken to assume normative values from literature or to monitor patients' progression because they should always be measured with the same stabilization method.

Stabilization, the participant's position, and the test angle should be taken into account to reach excellent reliability and, in this regard, examiner training is very important. Also, verbal encouragement and test instruction should be properly provided to reach the participant's maximum isometric strength. During the tests with the belt-stabilized method, the dynamometer should be well coupled to the participant's lower limb and the belt should be firm to prevent device displacement during the test and result alteration. The examiner should also be properly positioned to resist the force imposed by the participant. By assuming the use with the stabilization performed by the examiner, it should take into account that the examiner's hand grip strength and gender might also influence in the strength measurement of lower limb (Ieiri et al., 2015; Krause et al., 2014). If the participant's strength is greater than the examiner or a stable position cannot be achieved, than belt-stabilization method is advised.

This study has some limitations. The order of the exercises was not completely randomized, as we wanted to prevent many changes in the position and optimize the collection time. We also observed difficult control of the HHD stabilized with a belt in some movements, requiring better attention from the examiner to correct the dynamometer position. The study sample involved only young and healthy participants, limiting result generalization for the elderly population and subjects with a lower limb dysfunction, requiring future studies. The lower limb dominance and the gravity effects on the strength due to participant position were not considered which would provide a better understanding among the errors and agreement between both stabilization methods. However, it is also the first study to perform an agreement analysis between belt and examiner stabilization of HHD and to demonstrate the magnitude of the bias and to elucidate proportional bias between both methods. It consolidate that although both methods are reliable, they are not equivalent for knee and hip strength measurements.

5. Conclusions

The Lafayette® HHD presents excellent reliability for the evaluation of muscle groups of the hip and knee for methods stabilized by a belt or an examiner in young and healthy participants. However, lower error values and greater measures are achieved when the examiner stabilizes the HHD.

Acknowledgement

We acknowledge to São Paulo Research Foundation for the grant (2014/16120-8) that financially supported the development of this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ptsp.2019.04.011.

Conflicts of interest

The authors report no declarations of interest.

Ethical approval

The subjects received and signed an informed consent form approved by the Research Ethics Committee related to University of São Paulo (Case No. 15917/2014).

Funding

Third author received a grant from São Paulo Research Foundation to develop this research.

References

Barbosa, M. B., Montebelo, M. I. L., & Guirro, E. C. O. (2007). Determination of sensory perception and motor response thresholds in different phases of the menstrual cycle. *Revista Brasileira de Fisioterapia, 11*(6), 443–449.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*(2), 135–160.

Bohannon, R. W., Chu, J., & Portz, M. (2015). Measurement of hip extension strength with a portable device: Description, reliability and validity of a procedure. *Isokinetics and Exercise Science, 23*(4), 271–274.

Bohannon, R. W., Kindig, J., Sabo, G., Duni, A. E., & Cram, P. (2012). Isometric knee extension force measured using a handheld dynamometer with and without belt-stabilization. *Physiotherapy Theory and Practice, 28*(7), 562–567.

Bohannon, R. W., Pritchard, R. O., & Glenney, S. S. (2013). Portable belt-stabilized hand-held dynamometry set-up for measuring knee extension force. *Isokinetics and Exercise Science, 21*(4), 325–329.

Bonett, D. G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine, 21*(9), 1331–1335.

Chamorro, C., Armijo-Olivo, S., De La Fuente, C., Fuentes, J., & Javier Chirosa, L. (2017). Absolute reliability and concurrent validity of hand held dynamometry and isokinetic dynamometry in the hip, knee and ankle joint: Systematic review and meta-analysis. *Open Medicine (Poland), 12*(1), 359–375.

Fleiss, J. (1986). *Reliability of measurement. The design and analysis of clinical experiments*.

Fulcher, M. L., Hanna, C. M., & Raina Elley, C. (2010). Reliability of handheld dynamometry in assessment of hip strength in adult male football players. *Journal of Science and Medicine in Sport, 13*(1), 80–84.

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica, 25*(2), 141–151.

Halabchi, F., Mazaheri, R., & Seif-Barghi, T. (2013). Patellofemoral pain syndrome and modifiable intrinsic risk factors; how to assess and address? *Asian Journal of Sports Medicine, 4*(2), 85–100.

Hansen, E. M., McCartney, C. N., Sweeney, R. S., Palimenio, M. R., & Grindstaff, T. L. (2015). Hand-held dynamometer positioning impacts discomfort during quadriceps strength testing: A validity and reliability study. *International Journal of Sports Physical Therapy, 10*(1), 62–68.

Ieiri, A., Tushima, E., Ishida, K., Inoue, M., Kanno, T., & Masuda, T. (2015). Reliability of measurements of hip abduction strength obtained with a hand-held dynamometer. *Physiotherapy Theory and Practice, 31*(2), 146–152.

Jackson, S. M., Cheng, M. S., Smith, A. R., Jr., & Kolber, M. J. (2017). Intrarater reliability of hand held dynamometry in measuring lower extremity isometric strength using a portable stabilization device. *Musculoskeletal Science and Practice, 27*, 137–141.

Kelln, B. M., McKeon, P. O., Gontkof, L. M., & Hertel, J. (2008). Hand-held dynamometry: Reliability of lower extremity muscle testing in healthy, physically active,young adults. *Journal of Sport Rehabilitation, 17*(2), 160–170.

Kim, W. K., Kim, D., Seo, K. M., & Kang, S. H. (2014). Reliability and validity of isometric knee extensor strength test with hand-held dynamometer depending on its Fixation : A pilot study. *Annals of Rehabilitation Medicine, 38*(1), 84–93.

Kim, S.-G., & Lee, Y.-S. (2015). The intra- and inter-rater reliabilities of lower extremity muscle strength assessment of healthy adults using a hand held dynamometer. *Journal of Physical Therapy Science, 27*(6), 1799–1801.

Kollock, R. O., Onate, J. A., & Van Lunen, B. (2010). The reliability of portable fixed dynamometry during hip and knee strength assessments. *Journal of Athletic Training, 45*(4), 349–356.

Krause, D. A., Neuger, M. D., Lambert, K. A., Johnson, A. E., DeVinny, H. A., & Hollman, J. H. (2014). Effects of examiner strength on reliability of hip-strength testing using a handheld dynamometer. *Journal of Sport Rehabilitation, 23*(1), 56–64.

Martins, J., da Silva, J. R., da Silva, M.R.B., & Bevilaqua-Grossi, D. (2017). Reliability and validity of the belt-stabilized handheld dynamometer in hip- and knee-strength tests. *Journal of Athletic Training, 52*(6), 1062–6050–52.6.04.

Matsudo, S. M., Matsudo, V. R., Araújo, T., Andrade, D., Andrade, E., & Oliveira, L. (2002). Physical activity level of São Paulo state population: An analysis based on gender, age, socio-economic status, demographics and knowledge. *Revista Brasileira de Ciência e Movimento, 10*(4), 2002.

Mentiplay, B. F., Perraton, L. G., Bower, K. J., Adair, B., Pua, Y. H., Williams, G. P., et al. (2015). Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: A reliability and validity study. *PLoS One, 10*(10), 1–18.

Oliveira, I. O., Pilz, B., Santos Junior, R. L. G., Vasconcelos, R. A., Mello, W., & Grossi, D. B. (2018). Reference values and reliability for lumbopelvic strength and endurance in asymptomatic subjects. *Brazilian Journal of Physical Therapy, 22*(1), 33–41.

Prentice, W. E., & Kaminski, T. W. (2004). *Rehabilitation techniques for sports medicine and athletic training*.

Scott, D. A., Bond, E. Q., Sisto, S. A., & Nadler, S. F. (2004). The intra- and interrater reliability of hip muscle strength assessments using a handheld versus a portable dynamometer anchoring station. *Archives of Physical Medicine and Rehabilitation, 85*(4), 598–603.

Thorborg, K., Bandholm, T., & Hölmich, P. (2013). Hip and knee strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable. *Knee Surgery, Sports Traumatology, Arthroscopy, 21*(3), 550–555.

Toonstra, J., & Mattacola, C. G. (2013). Test retest reliability and validity of isometric knee flexion and extension measurement using 3 methods of assessing muscle strength. *Journal of Sport Rehabilitation*, (7), 1–5.

Wadsworth, C., Nielsen, D. H., Corcoran, D. S., Phillips, C. E., & Sannes, T. L. (1992). Interrater reliability of hand-held dynamometry: Effects of rater gender, body weight, and grip strength. *Journal of Orthopaedic & Sports Physical Therapy, 16*(2), 74–81.

Wikholm, J. B., & Bohannon, R. W. (1991). Hand-held dynamometer measurements: Tester strength makes a difference. *Journal of Orthopaedic & Sports Physical Therapy, 13*(4), 191–198.