Journal of Athletic Training 2017;52(9):809–819
doi: 10.4085/1062-6050-52.6.04

www.natajournals.org

Reliability and Validity of the Belt-Stabilized Handheld Dynamometer in Hip- and Knee-Strength Tests

Jaqueline Martins, PT; Janaina Rodrigues da Silva, PT; Marcelo Rodrigues Barbosa da Silva; Débora Bevilaqua-Grossi, PhD

Ribeirão Preto Medical School, University of São Paulo, Brazil

Context: The belt-stabilized handheld dynamometer (HHD) has been used to assess the strength of knee- and hip-muscle groups. However, few researchers have examined its reliability and validity for assessing the strength of these muscles.

Objective: To evaluate the intra-examiner reliability of the belt-stabilized HHD and its validity and agreement with the isokinetic dynamometer for assessing the strength of knee- and hip-muscle groups.

Design: Cross-sectional study.

Setting: University laboratory.

Patients or Other Participants: We evaluated 26 healthy participants (13 men, 13 women; age = 23.5 ± 2.8 years, height = 1.7 ± 0.1 m, mass = 68.6 ± 12.4 kg) in 2 sessions using the belt-stabilized HHD and an isokinetic dynamometer for maximum strength of the hip adductors, abductors, flexors, extensors, internal rotators, and external rotators and the knee flexors and extensors.

Main Outcome Measure(s): We used reliability values provided by the intraclass correlation coefficient (2,3), standard error of measurement (SEM and percentage SEM), and minimal detectable change; correlation values comparing the belt-stabilized HHD and the isokinetic instrument using the Pearson correlation coefficient (r); and the mean difference in values comparing the 2 instruments using the Bland-Altman method.

Results: The intrarater HHD reliability was excellent for most measurements (range = 0.80–0.96; SEM = 1.3–5.3 kilograms of force or 4.8–18.9 Nm, percentage SEM = 7.0%–22.0%, minimal detectable change = 3.6–18.8 kilograms of force or 13.2–52.4 Nm) and was moderate only for bilateral knee flexion and left hip internal rotation (intraclass correlation coefficient [2,3] = 0.62–0.66 and 0.70, respectively). Correlation with the isokinetic dynamometer was moderate to high (r = 0.60–0.90), but the absolute values did not demonstrate concordance between results using the Bland-Altman method.

Conclusions: The belt-stabilized HHD measurements were reliable, and although they did not agree with those from the isokinetic dynamometer, the values were correlated for the hip- and knee-muscle groups.

Key Words: muscle strength, lower extremity, reproducibility of results

Key Points

- The belt-stabilized hand-held dynamometer (HHD) displayed moderate to excellent intra-examiner reliability for measuring the strength of the knee- and hip-muscle groups.
- The belt-stabilized HHD was moderately to highly correlated with the isokinetic dynamometer for measuring the strength of the knee- and hip-muscle groups.
- The HHD is an alternative to the isokinetic dynamometer for evaluating knee- and hip-muscle strength.

Some pathologic musculoskeletal conditions of the lower extremity, such as patellofemoral pain syndrome (PFPS), knee pain after anterior cruciate ligament reconstruction, and knee osteoarthritis, are closely related to muscular deficits and imbalances in the knee and hip joints.[1–3] A strength deficit in the quadriceps, for example, can be a risk factor for PFPS,[3–5] is associated with knee pain in osteoarthritis,[3] and is a predictor of knee function and knee biomechanics after ACL rupture.[6,7] Imbalance or weakness of the hip muscles is a factor generally associated with and not a risk factor for PFPS,[3,8–10] and hip-abductor strength is associated with the results of performance-based functional tests in patients with knee arthroplasty[11] and osteoarthritis.[12] Therefore, assessing knee- and hip-muscle strength is important for a variety of pathologic lower extremity musculoskeletal conditions and is useful for rehabilitation programs aimed at restoring muscle strength, providing valuable information on the effectiveness and progress of treatment.[13–15] However, the evaluation method must be reliable and valid.

The isokinetic dynamometer and handheld dynamometer (HHD) are 2 devices available for assessing muscle strength.[15] The isokinetic dynamometer is considered the criterion standard for assessing muscle strength and performance[14,16] because it is a reference method for other instruments that measure strength. For measuring isometric torque, the isokinetic dynamometer has excellent intratrial and test-retest reliability (0.99–1.0) and can produce valid measures with a 1% coefficient of variation of the method error.[17] However, the device lacks portability, has a high cost, and requires extensive physical space and expert examiner training.[13,18,19] The HHD is used in clinical practice to assess isometric muscle strength as an alternative to the isokinetic dynamometer,[19] and several researchers have shown its considerable advantage in being portable, inexpensive, and easy to use.

The HHD has been tested while stabilized by an examiner and demonstrated reliable measurements of lower limb strength. Given that examiner strength is a known concern with the HHD, the belt-stabilized HHD protocol is potentially more reliable and valid, and researchers have recommended its use.[18,20–22]

The belt-stabilized HHD has shown excellent reliability in evaluating the strength of the knee extensors[13,15,23,24] and moderate to excellent reliability for the knee flexors.[15,25] For the hip, investigators have observed excellent reliability for the abductors,[25–27] adductors,[25] flexors,[25] and extensors.[25] However, most of the hip muscles were investigated in healthy athletes only by Thorborg et al,[25] and the rotator muscles were not studied at all. The belt-stabilized HHD has also been tested for validity against the isokinetic dynamometer.[13,24,28] The knee extensors were the most studied muscle group, and the HHD displayed moderate to high correlation (r range = 0.47–0.93)[13,24,28] except in the work by Kim et al,[24] who found a weak correlation (0.29) that was not different ($P > .05$) in the sitting position for the left knee extensor. Only Katoh et al[28] investigated the correlation between the HHD and the isokinetic dynamometer for the knee flexors and hip-muscle groups and observed a high correlation ($r = 0.88$) for the knee flexors and moderate to high correlation (r range = 0.52–0.86) for the hip muscles; however, the correlation was low ($r = 0.34$) for the abductor muscles.

Whereas some investigators have shown that the belt-stabilized HHD is reliable and valid for assessing strength of the knee muscles, few researchers have examined its reliability for the hip muscles, and none have reported on the rotator muscles. Katoh et al[28] were the only researchers to test its validity for the hip muscles, and they described a weak association with the isokinetic dynamometer values for the abductor muscles. The validity of the HHD has not been established for the hip; more studies are needed. Therefore, the purpose of our study was to evaluate the test-retest reliability of the belt-stabilized HHD and its validity and concordance compared with the isokinetic dynamometer for assessing hip- and knee-muscle strength.

METHODS

Study Design

The study was conducted at a university research laboratory and consisted of 2 assessments in a cross-sectional design.

Participants and Examiners

We selected 26 individuals (13 men, 13 women; age = 23.5 ± 2.8 years [range = 18–28 years], height = 1.7 ± 0.1 m, mass = 68.6 ± 12.4 kg) recruited by convenience from a university community. Volunteers were excluded if they reported pain, history of injury, orthopaedic surgery of the lower limbs in the 6 months before the study, or any neurologic disorder or if their physical activity was classified as *very active* by the International Physical Activity Questionnaire.[29]

The sample size was calculated on the basis of 2 observations per participant, an expected intraclass correlation coefficient (ICC) of 0.70, and an acceptable amplitude of the confidence interval (CI) for the ICC of 0.40, yielding 26 individuals.[30] All volunteers provided written informed consent, and the study was approved by the research ethics committee of our university (Case No. 15685/2013). Our methods followed the recommendations of Guidelines for Reporting Reliability and Agreement Studies.[31]

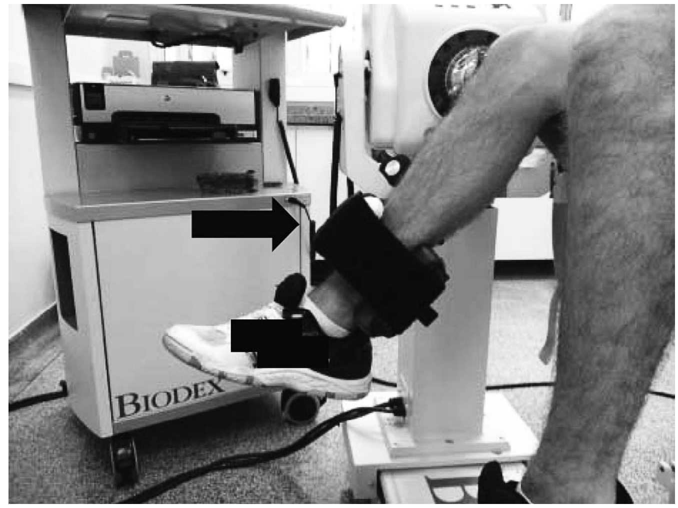

Figure 1. Hand-held dynamometer attached to the isokinetic dynamometer with hook-and-loop straps.

The main examiner (J.R.S.) has a bachelor's degree in physical therapy, which is achieved after 5 years of university educational training and requires diploma registration at our federal council of physical therapy. This examiner, who had 6 years of experience, conducted the strength tests. An undergraduate student in physical therapy (M.R.B.S.) read and recorded the values obtained with the dynamometers. Both examiners received a 30-hour training session on lower limb muscle-strength tests using the HHD and isokinetic dynamometer.

Procedures

Strength tests of bilateral knee extensors and flexors and hip internal and external rotators, extensors, flexors, adductors, and abductors were performed using an HHD (model 01163; Lafayette Instrument Company, Lafayette, IN) coupled distally to the lever arm of an isokinetic dynamometer (model Biodex 3; Biodex Medical Systems Inc, Shirley, NY; Figure 1). Hook-and-loop straps stabilized the instrument, ensuring the simultaneous collection of data and preventing examiner influence (Figure 1).

The volunteers were positioned in the isokinetic dynamometer chair with their torso and hips stabilized with hook-and-loop straps, and all testing positions were consistent with the manufacturer's instructions. The knee flexors and extensors and hip internal and external rotators were tested in the sitting position with the knee and hip flexed to 90° and the HHD positioned at the ankle 2 cm above the malleoli.[13,15] We tested the hip adductors and abductors in the lateral decubitus position with the hip abducted to 30° and in neutral rotation. The hip flexors and extensors were tested in the supine position with the hip flexed to 30° and in neutral rotation. These 4 tests were performed with the HHD placed at the distal thigh 2 cm above the femoral epicondyle (Figure 2).[28] The order of the

Figure 2. Belt-stabilized hand-held dynamometer and isokinetic dynamometer tests. A, Knee flexion and extension. B, Hip internal rotation and external rotation. C, Hip adduction and abduction. D, Hip flexion and extension.

joints (knee and hip) was randomized, the order of the muscle groups was defined by convenience, and the right limb was tested before the left limb. When the tests began with the knee joint, the sequence was knee flexion and extension followed by hip internal rotation, external rotation, extension, flexion, adduction, and abduction. When the tests began with the hip joint, the sequence was hip abduction, adduction, flexion, extension, external rotation, and internal rotation followed by knee flexion and extension. The order of muscle groups was established to prioritize patient positioning (sitting and lying) to avoid constant changes of positions, minimizing the collection time and reducing participant fatigue.

We initially collected anthropometric data for height, body mass, and length of the thigh and leg (measured from the greater trochanter to the lateral epicondyle of the femur and from the lateral epicondyle of the femur to the lateral malleolus of the tibia). Volunteers performed a warm-up on a bicycle with a light load for 10 minutes and then 3 submaximal isometric contractions for each test position on an isokinetic dynamometer to become familiar with the device. For each muscle group, they performed 3 maximal voluntary isometric contractions for 5 seconds, with 60 seconds of rest between repetitions and 5 minutes of rest between tests. Visual feedback was provided by the graphed torque versus time shown on the isokinetic dynamometer monitor. However, visual feedback was not possible for tests carried out in the supine position due to the inability of participants to view the monitor. We instructed participants to use maximal voluntary effort and provided oral encouragement to maintain the strength initially displayed during the test. Under the same conditions as the first evaluation, the tests were repeated in the same order by the same examiner about 7 to 10 days after the first session, and the average interval was 8.3 days (Figure 2).

Statistical Analysis

To analyze the HHD test-retest reliability, we used the ICC (2,3) from the average values of 3 repetitions performed on each strength test in the first and second sessions. The ICC was associated with a 95% CI, and reliability was classified as *poor* (ICC $<$ 0.40), *moderate* (ICC $=$ 0.40–0.75), or *excellent* (ICC $>$ 0.75).[32] We performed *t* tests for the mean values obtained from the

HHD in kilograms of force (kgf) between the first and second sessions to verify the variability of the strength values and set the α level at .05. The standard error of measurement ($SEM = SD\sqrt{1 - ICC}$) was calculated for the HHD, where *SD* indicated the standard deviation and the minimal detectable change ($MDC = SEM \times 1.96 \times \sqrt{2}$) was associated with a 95% CI.[33] The reliability value took into account the strength values in kilograms of force, and the error measurements were expressed in kilograms of force and newton · meters. The SEM was also expressed as a percentage of average strength (%SEM) for values in kilograms of force.

We analyzed the validity of the HHD using the Pearson product moment correlation coefficient (*r*), which verified the degree of correlation between the torques (measured in newton · meters) from the HHD and the isokinetic dynamometer, obtained in the second session to ensure the volunteers had greater familiarity with the tools and tests. The correlation was classified as *high* ($r \geq 0.70$), *moderate* ($0.70 > r \geq 0.40$), or *low* ($r < 0.40$),[34] and the α level was set at .05.

The Bland-Altman method was used to analyze the agreement between the measurements of the HHD and the isokinetic dynamometer, graphically displaying the average difference (isokinetic dynamometer torque minus HHD torque) in measures between devices and the 95% limits of agreement (LOAs).[35,36] The 95% LOAs were calculated as LOA = bias ± 1.96 SD, with *bias* indicating the mean difference between measurement methods. Linear regression analyses were conducted to verify the relationship between the difference and the magnitude of the measurement, thereby providing insights into proportional bias.[32] Whereas a logarithmic transformation of the data is recommended to control for increasing differences between measures as the magnitude of measure increases, we used the raw data because they are better interpreted clinically.[35,36] To compare the HHD measurements with the isokinetic dynamometer, we converted HHD strength data obtained in kilograms of force units into newtons and multiplied by the limb length, thereby obtaining the torque in newton · meters. For the HHD values in newton · meters that were used in the error measurements, correlation, and agreement analyses, the limb length was reduced by 2 cm due to the pad position in relation to anatomic landmarks. Comparisons were established for the average torque (Nm) of 3 repetitions performed in each test during the second session.

We used SPSS (version 17; SPSS Inc, Chicago, IL) for all analyses and to construct the Bland-Altman graphs.

RESULTS

Twenty-six participants attended the first and second sessions. The mean force and torque generated by each muscle group measured by the HHD in the first and second sessions and the mean torque measured on the isokinetic dynamometer in the second session are presented in Table 1. One device did not consistently register greater strength than the other: the strength data were higher for the HHD than for the isokinetic dynamometer for most muscle groups but lower for the knee flexors and extensors and hip internal rotators. The *t* tests showed no difference (t_{25} range $= -1.02$ to 1.47, $P < .05$) between the HHD mean values in

Table 1. Belt-Stabilized Handheld and Isokinetic Dynamometer Force Measurements by Muscle Group (N = 26; Mean ± SD)

	Belt-Stabilized Hand-Held Dynamometer, kgf (Nm)				Isokinetic Dynamometer, Nm	
	Right Limb		Left Limb		Right Limb	Left Limb
	Session					
Muscle Group	First	Second	First	Second	Second	Second
Knee extensors	43.3 ± 17.6 (163.8 ± 76.0)	43.7 ± 17.8 (167.3 ± 79.7)	41.7 ± 18.2 (156.6 ± 73.9)	41.6 ± 18.5 (156.4 ± 78.0)	192.0 ± 77.1	176.1 ± 73.2
Knee flexors	21.6 ± 8.3 (81.7 ± 35.9)	22.1 ± 7.7 (84.7 ± 29.4)	20.3 ± 6.4 (76.3 ± 28.4)	20.5 ± 7.3 (77.5 ± 32.4)	92.3 ± 31.8	93.8 ± 33.1
Hip internal rotators	10.2 ± 4.6 (38.0 ± 17.8)	10.0 ± 4.3 (36.2 ± 15.6)	8.8 ± 2.7 (32.9 ± 11.4)	9.0 ± 3.2 (34.7 ± 15.7)	49.6 ± 15.3	46.8 ± 14.9
Hip external rotators	9.5 ± 4.2 (35.8 ± 18.7)	9.2 ± 4.2 (33.9 ± 17.6)	7.8 ± 3.9 (29.4 ± 15.7)	7.9 ± 3.6 (30.2 ± 13.0)	15.9 ± 12.2	15.2 ± 10.3
Hip extensors	39.3 ± 13.6 (149.5 ± 52.7)	40.1 ± 14.9 (156.2 ± 65.9)	39.9 ± 17.1 (152.8 ± 70.4)	39.4 ± 16.8 (149.5 ± 69.5)	139.1 ± 55.5	132.9 ± 46.2
Hip flexors	29.6 ± 8.8 (114.0 ± 39.1)	29.2 ± 8.6 (110.5 ± 36.3)	28.4 ± 9.0 (109.3 ± 39.1)	28.3 ± 9.2 (108.4 ± 41.8)	89.6 ± 38.9	87.0 ± 35.7
Hip adductors	30.8 ± 10.1 (119.7 ± 46.9)	29.9 ± 10.5 (112.3 ± 50.1)	30.7 ± 11.2 (118.6 ± 50.3)	30.0 ± 10.6 (113.3 ± 44.0)	93.5 ± 37.7	95.6 ± 29.7
Hip abductors	25.6 ± 7.8 (97.9 ± 32.8)	25.6 ± 7.1 (98.2 ± 29.0)	24.8 ± 8.0 (94.8 ± 32.2)	24.3 ± 7.6 (91.1 ± 30.5)	82.2 ± 26.2	85.0 ± 26.4

Abbreviations: kgf, kilograms of force; Nm, newton · meters.

Table 2. Reliability of Belt-Stabilized Handheld Dynamometer Measurements (N = 26)

Muscle Group	Right Limb: Intraclass Correlation Coefficient (2,3) (95% CI)	Right Limb: Standard Error of Measurement kgf (Nm)	Right Limb: Standard Error of Measurement %	Right Limb: Minimal Detectable Change, kgf (Nm)	Left Limb: Intraclass Correlation Coefficient (2,3) (95% CI)	Left Limb: Standard Error of Measurement kgf (Nm)	Left Limb: Standard Error of Measurement %	Left Limb: Minimal Detectable Change, kgf (Nm)
Knee extensors	0.91 (0.82, 0.96)	5.3 (18.9)	12.0	14.8 (52.4)	0.93 (0.84, 0.96)	4.8 (18.4)	12.0	13.5 (51.1)
Knee flexors	0.66 (0.26, 0.85)	4.4 (16.3)	20.0	12.4 (45.1)	0.62 (0.14, 0.83)	4.4 (17.6)	22.0	12.4 (48.8)
Hip internal rotators	0.80 (0.57, 0.91)	1.9 (6.8)	19.0	5.3 (18.9)	0.70 (0.34, 0.86)	1.7 (6.9)	20.0	4.9 (19.2)
Hip external rotators	0.90 (0.79, 0.95)	1.3 (4.8)	14.0	3.6 (13.2)	0.80 (0.55, 0.91)	1.5 (5.9)	20.0	4.4 (16.3)
Hip extensors	0.91 (0.80, 0.96)	4.4 (17.7)	11.0	12.4 (49.2)	0.94 (0.88, 0.97)	4.1 (15.5)	10.0	11.4 (42.9)
Hip flexors	0.90 (0.79, 0.95)	2.7 (9.9)	9.0	7.5 (27.4)	0.96 (0.91, 0.98)	1.8 (6.9)	7.0	5.1 (19.3)
Hip adductors	0.88 (0.74, 0.94)	3.6 (13.6)	12.0	10.0 (37.8)	0.90 (0.78, 0.95)	3.3 (13.3)	11.0	9.2 (36.7)
Hip abductors	0.81 (0.58, 0.91)	3.1 (11.9)	12.0	8.6 (33.0)	0.87 (0.72, 0.94)	2.7 (9.8)	11.0	7.5 (27.3)

Abbreviations: CI, confidence interval; kgf, kilograms of force; Nm, newton · meters.

kilograms of force between the first and second sessions (Table 1).

Reliability

The ICC values of the HHD ranged from 0.62 to 0.96, indicating excellent test-retest reliability for most muscle groups except the bilateral knee flexors and internal-rotator muscles of the left hip; for these muscles, reliability was moderate and the CI was poor to excellent, ranging from 0.14 to 0.85 for the knee flexors and 0.34 to 0.86 for the left hip internal-rotator muscles. The SEM (range = 4.4–5.3 kgf [16.3–18.9 Nm]) and MDC (12.4–14.8 kgf [45.1–52.4 Nm]) values were greater in the knee-muscle groups. For the hip-muscle groups, SEM values ranged from 1.3 to 4.4 kgf (4.8 to 17.7 Nm), and MDC values ranged from 3.6 to 12.4 kgf (13.2 to 49.2 Nm). The %SEM values ranged from 12.0% to 22.0% for the knee-muscle groups and from 7.0% to 20.0% for the hip-muscle groups (Table 2).

Validity

We observed a high correlation between the HHD and the isokinetic dynamometer bilaterally for the knee extensors and hip adductors (r range = 0.78–0.90). Values for the hip external rotators were moderately correlated for both limbs (r range = 0.60–0.63). For the other muscles, the correlation ranged from moderate to high among limbs (r range = 0.61–0.84; Table 3).

Agreement

Bland-Altman plots showed that the belt-stabilized HHD tended to overestimate the strength of most hip-muscle groups in relation to the isokinetic dynamometer, mainly for hip flexion, external rotation, adduction, and extension at values higher than 140 Nm and 175 Nm, respectively, suggesting a possible systematic bias. The HHD magnitude was generally lower than the isokinetic dynamometer magnitude only for hip internal rotation. For the knee-muscle groups, isokinetic values were generally higher than HHD values, but only random error was found. Furthermore, the difference tended to increase as the mean increased for some hip-muscle groups, such as abduction (mainly for values greater than 100 Nm), and we observed a difference for the slope of the regression line for right (P = .004) and left (P = .003) hip adduction, left hip extension (P = .02), and right hip external rotation (P = .03). This relationship was not observed for the knee-muscle groups. The average difference between the HHD and isokinetic dynamometer ranged from 7.5 to 24.7 Nm for the knee-muscle groups, and the limits of agreement ranged from –76.4 to 125.8 Nm (Figure 3). The data for the hip-muscle groups are presented in Figures 4 and 5. We found no agreement between the instruments for either muscle group because all average differences (biases) were significant in both limbs and unilaterally for the knee flexors and hip extensors and abductors ($P < .05$); also, the limits of agreement showed a large difference between the devices.

DISCUSSION

Measurements from the belt-stabilized HHD were reliable, valid, and correlated with those from the isokinetic dynamometer for measuring the strength of the knee and hip muscles. However, the Bland-Altman plots showed no agreement between measurements made with the belt-stabilized HHD and the isokinetic dynamometer.

The belt-stabilized HHD demonstrated excellent test-retest reliability between sessions (ICC range = 0.91–0.93) for assessing knee-extensor strength; this finding agreed with the findings of other studies in which the reliability values ranged from 0.76 to 0.98 in young, healthy adults.[13,15,23,24] We observed moderate reliability (ICC range = 0.62–0.66) for knee-flexor strength, in agreement with the study by Toonstra and Mattacola,[15] who reported an ICC of 0.49 for the belt-stabilized HHD. The test-retest reliability of the belt-stabilized HHD was excellent for all hip-muscle groups (ICC = 0.80–0.96) and was moderate only for the right hip internal rotators (ICC = 0.70). Our results agreed with those reported by Kramer et al[26] (ICC

Table 3. Pearson Correlation Coefficient (r) Between Belt-Stabilized Handheld Dynamometer and Isokinetic Dynamometer Measurements (N = 26)

Muscle Group	Limb: Right	Limb: Left
Knee extensors	0.78	0.87
Knee flexors	0.74	0.69
Hip internal rotators	0.69	0.74
Hip external rotators	0.63	0.60
Hip extensors	0.84	0.61
Hip flexors	0.68	0.71
Hip adductors	0.90	0.81
Hip abductors	0.62	0.72

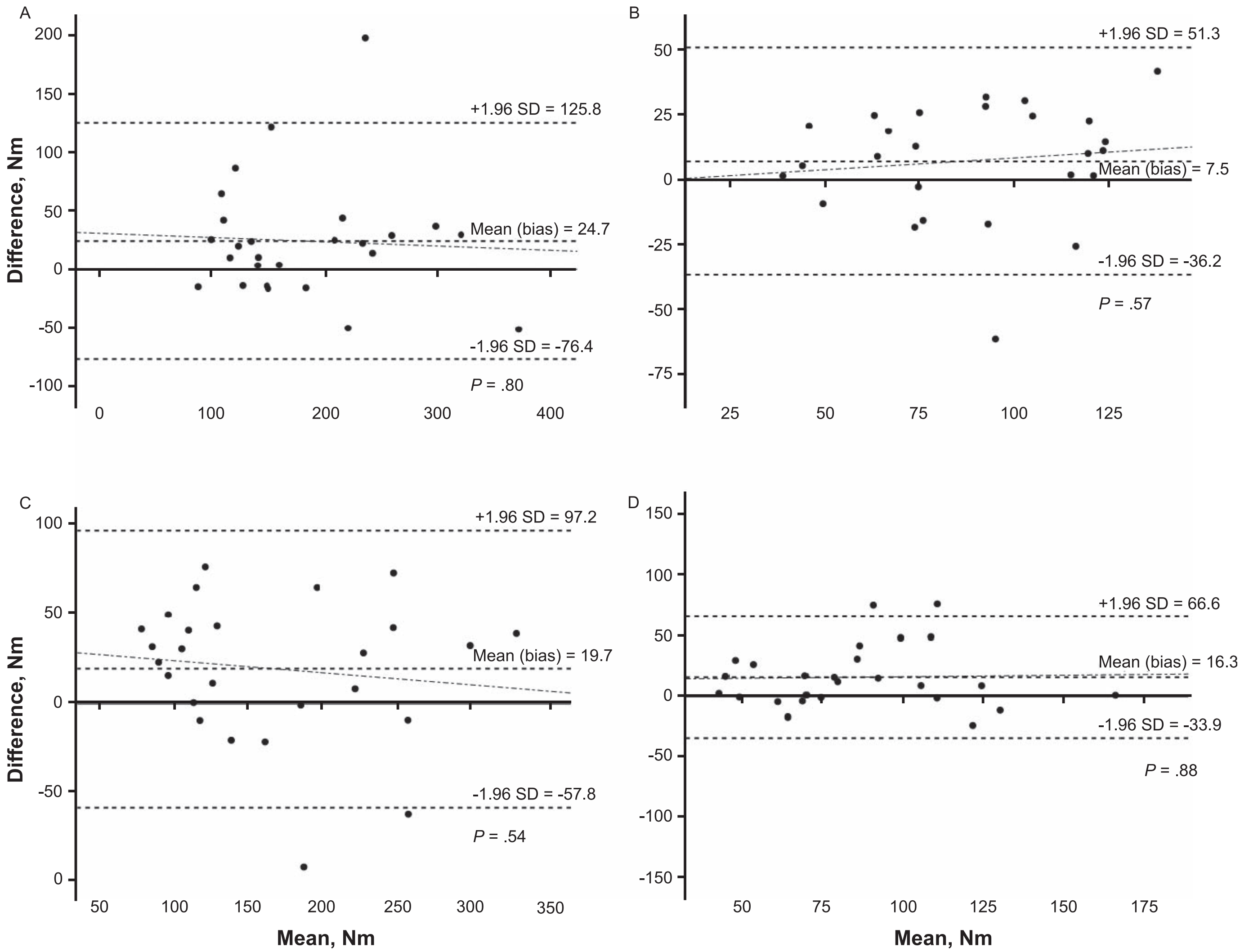

Figure 3. Bland-Altman plots comparing the belt-stabilized hand-held dynamometer and isokinetic dynamometer measurements in assessing torque of the knee extensors and flexors. The *P* value is for the slope of the regression line. A, Right knee extension (limits of agreement). B, Right knee flexion. C, Left knee extension. D, Left knee flexion. Abbreviation: SD, standard deviation.

range = 0.80–0.97) and Ieiri et al[27] (ICC range = 0.98–0.99) for the hip abductors. In their investigation of the reliability of other hip muscles, Thorborg et al[25] analyzed interrater reliability and found excellent values (ICC range = 0.76–0.85). No researchers have examined the reliability of the belt-stabilized HHD to assess the hip rotators. The reliability of the HHD without belt stabilization for hip-muscle groups has ranged from 0.82 to 0.97,[20–22,24] which is in accordance with the values we presented. Reliability values reported in other studies include 0.95 to 0.96 for the hip external rotators,[21] 0.82 to 0.84 for the hip extensors,[22] 0.83 to 0.92 for the hip flexors,[22] and 0.85 to 0.93 for the hip abductors.[21,22] No authors examining the reliability of an HHD without belt stabilization for the hip internal rotators have reported lower reliability values than those we found.

The strength values obtained from the HHD should be interpreted by considering the SEM and MDC, which assist in clinical decision making. We reported the SEMs associated with test sessions, which varied from 4.4 to 5.3 kgf (12.0%–22.0% of the force generated) or 16.3 to 18.9 Nm for the knee-muscle groups and from 1.3 to 4.4 kgf (7.0%–20.0% of the force generated) or 4.8 to 17.7 Nm for the hip-muscle groups. For the MDC, a real change in the strength of the knee flexors and extensors is expected at values exceeding 12.4 kgf or 48.8 Nm and 14.8 kgf or 52.4 Nm, respectively. A real change in the strength of the hip-muscle groups is expected to exceed the muscle-specific values of 4.4 kgf or 16.3 Nm for the external rotators and 12.4 kgf or 49.2 Nm for the extensors.

The SEM and MDC values in newton · meters for the knee extensors were higher than the values reported by other authors using the belt-stabilized HHD,[13,15,24] and these higher values may have resulted from the large SDs of our data. In addition, the test position or limb lengths used to obtain the torque measure were different in our study, making comparisons difficult. Kim et al[24] reported an SEM range from 1.96 Nm to 2.86 Nm for the knee extensors when participants were positioned differently than in our study. Hansen et al[13] found an SEM of 5.4 Nm and an MDC of 15.1 Nm, but the HHD was positioned 5 cm above the lateral malleolus. Finally, Toonstra and Mattacola[15] reported an SEM of 0.18 Nm and an MDC of 0.50 Nm. Some researchers have shown lower values of SEM (range = 0.7–

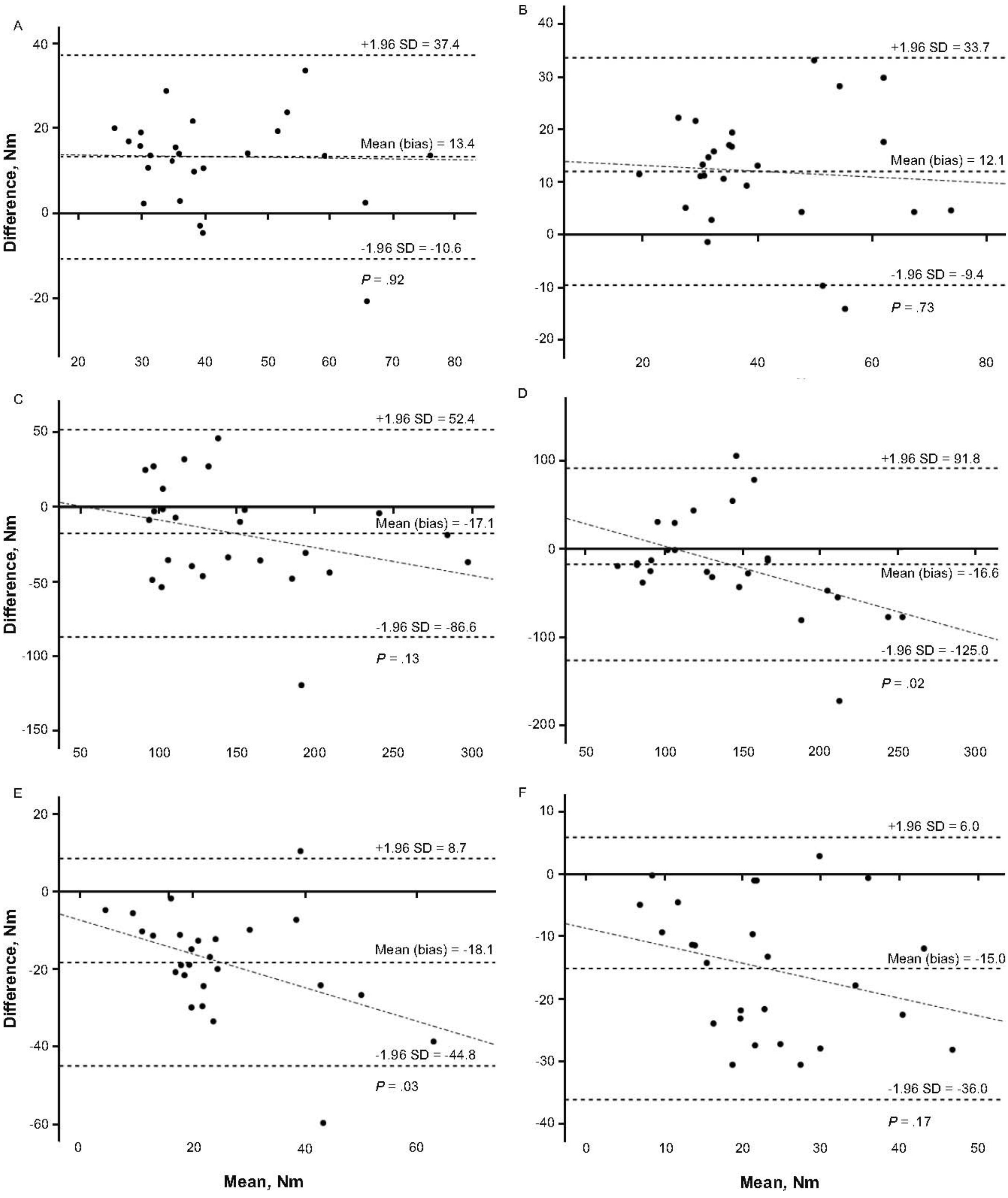

Figure 4. Bland-Altman plot comparing the belt-stabilized hand-held dynamometer and isokinetic dynamometer measurements in assessing torque of hip internal and external rotators and hip extensors. The *P* value is for the slope of the regression line. A, Right hip internal rotation. B, Left hip internal rotation. C, Right hip extension. D, Left hip extension. E, Right hip external rotation. F, Left hip external rotation. Abbreviation: SD, standard deviation.

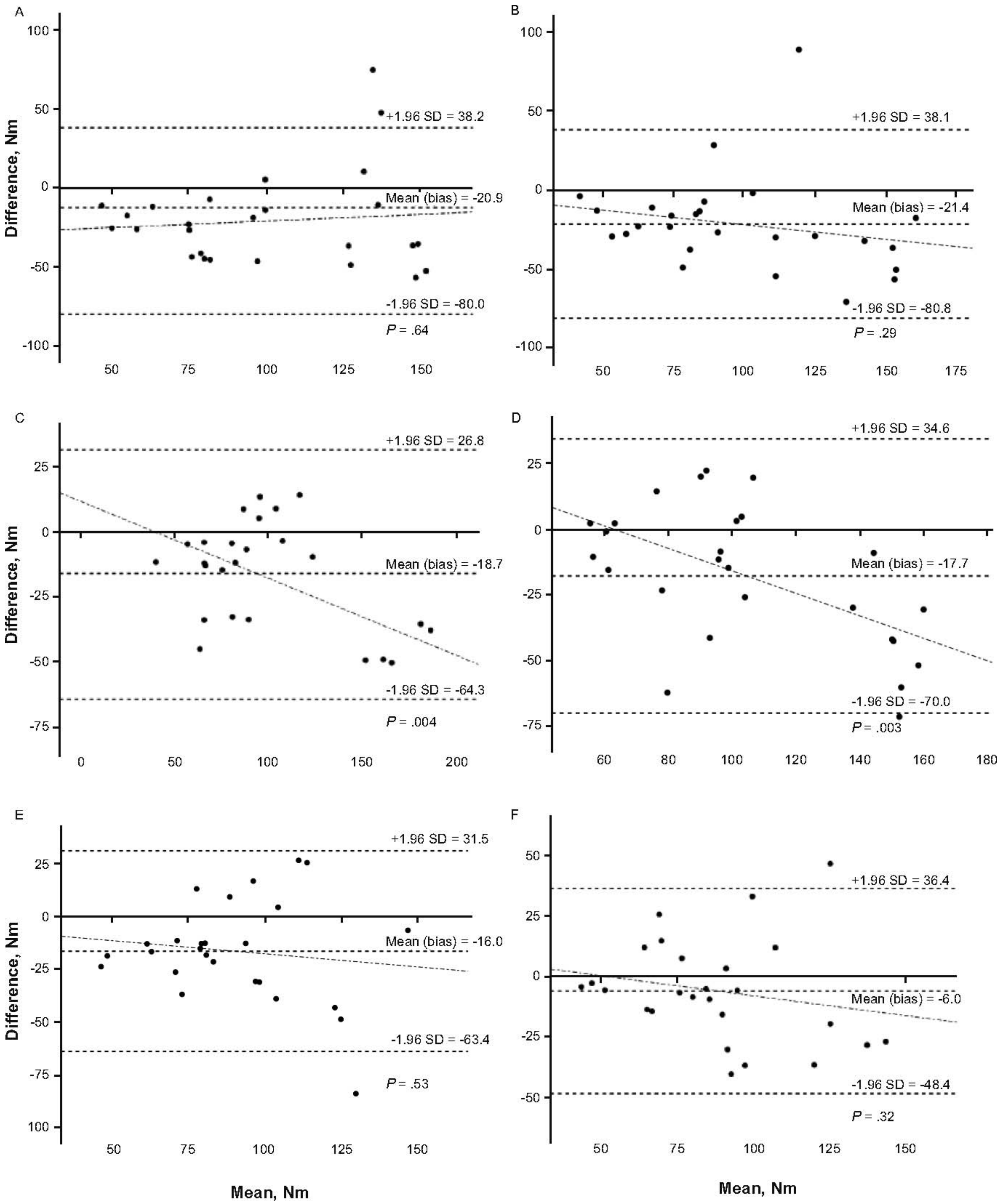

Figure 5. Bland-Altman plot comparing the belt-stabilized hand-held dynamometer and isokinetic dynamometer measurements in assessing torque of hip flexors, adductors, and abductors. The *P* value is for the slope of the regression line. A, Right hip flexion. B, Left hip flexion. C, Right hip adduction. D, Left hip adduction. E, Right hip abduction. F, Left hip abduction. Abbreviation: SD, standard deviation.

2.9 kgf or 7.3–28.9 N) and MDC (range = 1.7–8.2 kgf or 16.6–80.1 N) for the hip muscles.[26,27] However, in both studies, the investigators[26,27] analyzed the hip muscles using the HHD Power Track II with the participants in a supine position that was different from our test position, and Thorborg et al[25] assessed only healthy athletes. The SEM and MDC of the hip rotators were not analyzed using the belt-stabilized HHD. Our %SEM data indicated that the HHD generated more accurate values for the knee extensors and the hip extensors, flexors, adductors, and abductors. However, the HHD seemed to be less accurate for the knee flexors and internal and external hip rotators, which may indicate a greater need for test familiarization.

We found a high correlation (r range = 0.78–0.87) between the belt-stabilized HHD and the isokinetic dynamometer for knee-extensor strength, in agreement with Hansen et al[13] ($r = 0.93$) and Katoh et al[28] ($r = 0.75$). However, Kim et al[24] reported a low to moderate correlation (r range = 0.29–0.47) for the knee extensors evaluated in a sitting position using a belt-stabilized HHD. These lower correlations[24] probably occurred because the pelvis and trunk were not stabilized in the HHD test, allowing hip elevation and compromising the stability of muscular contractions, which did not occur in the isokinetic dynamometer test because the hip and torso were stabilized. Therefore, the stabilization belts used in our study adequately provided valid measures of knee-extensor strength.

The validity of the belt-stabilized HHD for the knee flexor and hip-muscle groups was analyzed only by Katoh et al,[28] who observed moderate to high correlations (r range = 0.52–0.88) similar to our observations (r = 0.60–0.90). We showed higher correlation values for the hip adductors (r range = 0.81–0.90) and abductors (r range = 0.62–0.72), but Katoh et al[28] reported moderate ($r = 0.52$) and low ($r = 0.34$) correlations, respectively. This divergence possibly occurred because the authors[28] used the supine test position and did not test some factors that could influence the force generated by the abductors and adductors, such as the pelvic and trunk stabilizers. Therefore, stabilizing the pelvis and trunk in the lateral decubitus position is considered more appropriate for increasing the validity of the HHD in these strength tests.

The Bland-Altman approach indicated a tendency toward systematic bias for some hip-muscle groups because the belt-stabilized HHD overestimated strength in almost all individuals for flexion, external rotation, adduction, and extension. In addition, the HHD underestimated the strength of hip internal rotation. For the knee muscles, the HHD generally underestimated the strength, and the difference was due to random error. We attributed random error to equipment variability or to differences in the variables measured by each dynamometer because strength was registered simultaneously to reduce the influence of the individual performance or observer variations. In addition, the expected difference between the HHD and the isokinetic dynamometer varied depending on the magnitude of torque for hip abduction, adduction, and external rotation and left hip extension. For most hip-muscle groups, the limits of agreement can be larger than necessary for small torque and narrower than they could be for large torque.[35] Therefore, despite the moderate to high correlations, the Bland-Altman method showed no agreement between the belt-stabilized HHD and the isokinetic dynamometer torque measurements because all muscle groups demonstrated significance for the average differences (biases), including systematic or proportional bias and random error, resulting in large limits of agreement. For example, the limits of agreement for the right knee extensors indicated that, when evaluating a patient, the clinician can expect the HHD values to be 76.4 Nm more or 125.8 Nm less than the isokinetic dynamometer values. Whereas the average strength of the knee extensors was 192 Nm for the isokinetic dynamometer, the HHD values can be 40% higher or 66% lower than the values of the isokinetic dynamometer, and these values have a clinical effect. Similar differences were observed between the HHD and the isokinetic dynamometer values in all muscle groups, with limits of agreement ranging from 20% for the hip internal rotators to 237% for the hip external rotators. When conducting this analysis, Hansen et al[13] measured only the strength of the knee extensors and concluded that agreement existed between the 2 methods. The difference between the 2 studies is probably due to the threshold for agreement that was based on clinical judgment.[35] The disagreement between the HHD and isokinetic dynamometer was expected because they originally measured different variables: the HHD measures strength in kilograms of force and the isokinetic dynamometer measures torque in newton · meters. Therefore, a simple manual conversion of HHD values from kilograms of force to newton · meters may not correspond with the automatic conversion by the isokinetic dynamometer. This measurement conversion is not enough to ensure agreement between the HHD and isokinetic dynamometer.

We showed that the absolute values generated by a manual dynamometer should not be interpreted as equivalent to those obtained from an isokinetic dynamometer and vice versa. Therefore, the HHD measurements are not a valid representation of the isokinetic measurements. However, the manual dynamometer can be considered a valid instrument, given that it measures what it purports to measure and shows a strong correlation with the criterion standard.

Our study had some limitations. The strength measures of the HHD were obtained with the HHD coupled to the isokinetic dynamometer. Despite this attempt to avoid the bias of nonsimultaneous collection of strength measures, this protocol does not reflect daily clinical practice. The results should be considered valid for testing with a belt-stabilized HHD only. They cannot be extrapolated to individuals in other age groups or with knee or hip disorders; these populations may present factors intrinsic to the individual, such as physical and mental health concerns, that interfere with reliability, validity, and even the ability to perform the test. In addition, the test-retest reliability for the bilateral knee flexors and left internal hip rotators was associated with a CI that included poor reliability values, probably due to the dominance or necessity of body stabilization.

Finally, our study had strengths. We confirmed the reliability and validity of the belt-stabilized HHD for all knee- and hip-muscle groups. The study was an innovative reliability analysis of the rotator muscles of the hip and analysis of the SEM, MDC, and agreement measurements for all knee- and hip-muscle groups. The HHD and isokinetic dynamometer measurements were recorded during the same contraction, which validates the correlation and agreement results, as well as the carefully developed and described methods, which were based on the Guidelines for Reporting Reliability and Agreement Studies[31]; this enables reproduction of the tests in the clinical setting.

CONCLUSIONS

The HHD is a reliable and valid method for evaluating the strength of the knee- and hip-muscle groups in a healthy, young adult population and may be used in clinical practice. However, its strength assessments do not agree with the force measurements of the isokinetic dynamometer.

ACKNOWLEDGMENTS

We thank the State of São Paulo Research Foundation (Process No. 2013/25755-4) for financial support.

REFERENCES

1. Muraki S, Akune T, Teraguchi M, et al. Quadriceps muscle strength, radiographic knee osteoarthritis and knee pain: the ROAD study. *BMC Musculoskelet Disord*. 2015;16:305.
2. Gadea F, Monnot D, Quelard B, et al. Knee pain after anterior cruciate ligament reconstruction: evaluation of a rehabilitation protocol. *Eur J Orthop Surg Traumatol*. 2014;24(5):789–795.
3. Lankhorst NE, Bierma-Zeinstra SMA, van Middelkoop M. Risk factors for patellofemoral pain syndrome: a systematic review. *J Orthop Sports Phys Ther*. 2012;42(2):81–94.
4. Pappas E, Wong-Tom WM. Prospective predictors of patellofemoral pain syndrome: a systematic review with meta-analysis. *Sports Health*. 2012;4(2):115–120.
5. Boling MC, Padua DA, Marshall SW, Guskewicz K, Pyne S, Beutler A. A prospective investigation of biomechanical risk factors for patellofemoral pain syndrome: the Joint Undertaking to Monitor and Prevent ACL Injury (JUMP-ACL) cohort. *Am J Sports Med*. 2009; 37(11):2108–2116.
6. Kline PW, Johnson DL, Ireland ML, Noehren B. Clinical predictors of knee mechanics at return to sport following ACL reconstruction. *Med Sci Sports Exerc*. 2015;48(5):790–795.
7. Eitzen I, Holm I, Risberg MA. Preoperative quadriceps strength is a significant predictor of knee function two years after anterior cruciate ligament reconstruction. *Br J Sports Med*. 2009;43(5):371–376.
8. Rathleff MS, Rathleff CR, Crossley KM, Barton CJ. Is hip strength a risk factor for patellofemoral pain? A systematic review and meta-analysis. *Br J Sports Med*. 2014;48(14):1088.
9. Magalhães E, Silva AP, Sacramento SN, Martin RL, Fukuda TY. Isometric strength ratios of the hip musculature in females with patellofemoral pain: a comparison to pain-free controls. *J Strength Cond Res*. 2013;27(8):2165–2170.
10. Thijs Y, Pattyn E, Van Tiggelen D, Rombaut L, Witvrouw E. Is hip muscle weakness a predisposing factor for patellofemoral pain in female novice runners? A prospective study. *Am J Sports Med*. 2011; 39(9):1877–1882.
11. Alnahdi AH, Zeni JA, Snyder-Mackler L. Hip abductor strength reliability and association with physical function after unilateral total knee arthroplasty: a cross-sectional study. *Phys Ther*. 2014;94(8): 1154–1162.
12. Tevald MA, Murray A, Luc BA, Lai K, Sohn D, Pietrosimone B. Hip abductor strength in people with knee osteoarthritis: a cross-sectional study of reliability and association with function. *Knee*. 2016;23(1): 57–62.
13. Hansen EM, McCartney CN, Sweeney RS, Palimenio MR, Grindstaff TL. Hand-held dynamometer positioning impacts discomfort during quadriceps strength testing: a validity and reliability study. *Int J Sports Phys Ther*. 2015;10(1):62–68.
14. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PM R*. 2011;3(5):472–479.
15. Toonstra J, Mattacola CG. Test–retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. *J Sport Rehabil*. https://www.ncbi.nlm.nih.gov/labs/articles/22951307/. Published February 18, 2013. Accessed April 11, 2017.
16. Roy MA, Doherty TJ. Reliability of hand-held dynamometry in assessment of knee extensor strength after hip fracture. *Am J Phys Med Rehabil*. 2004;83(11):813–818.
17. Drouin JM, Valovich-McLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex System 3 Pro isokinetic dynamometer velocity, torque and position measurements. *Eur J Appl Physiol*. 2004;91(1):22–29.
18. Martin HJ, Yule V, Syddall HE, Dennison EM, Cooper C, Aihie Sayer A. Is hand-held dynamometry useful for the measurement of quadriceps strength in older people? A comparison with the gold standard Biodex dynamometry. *Gerontology*. 2006;52(3):154–159.
19. Li RC, Jasiewicz JM, Middleton J, et al. The development, validity, and reliability of a manual muscle testing device with integrated limb position sensors. *Arch Phys Med Rehabil*. 2006;87(3):411–417.
20. Kim SG, Lee YS. The intra- and inter-rater reliabilities of lower extremity muscle strength assessment of healthy adults using a hand held dynamometer. *J Phys Ther Sci*. 2015;27(6):1799–1801.
21. Krause DA, Neuger MD, Lambert KA, Johnson AE, DeVinny HA, Hollman JH. Effects of examiner strength on reliability of hip-strength testing using a handheld dynamometer. *J Sport Rehabil*. 2014;23(1):56–64.
22. Arnold CM, Warkentin KD, Chilibeck PD, Magnus CR. The reliability and validity of handheld dynamometry for the measurement of lower-extremity muscle strength in older adults. *J Strength Cond Res*. 2010;24(3):815–824.
23. Katoh M. Reliability of isometric knee extension muscle strength measurements made by a hand-held dynamometer and a belt: a comparison of two types of device. *J Phys Ther Sci*. 2015;27(3):851–854.
24. Kim WK, Kim DK, Seo KM, Kang SH. Reliability and validity of isometric knee extensor strength test with hand-held dynamometer depending on its fixation: a pilot study. *Ann Rehabil Med*. 2014; 38(1):84–93.
25. Thorborg K, Bandholm T, Hölmich P. Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable. *Knee Surg Sports Traumatol Arthrosc*. 2013;21(3):550–555.
26. Kramer JF, Vaz MD, Vandervoort AA. Reliability of isometric hip abductor torques during examiner- and belt-resisted tests. *J Gerontol*. 1991;46(2):M47–M51.
27. Ieiri A, Tushima E, Ishida K, Inoue M, Kanno T, Masuda T. Reliability of measurements of hip abduction strength obtained with a hand-held dynamometer. *Physiother Theory Pract*. 2015;31(2): 146–152.
28. Katoh M, Hiiragi Y, Uchida M. Validity of isometric muscle strength measurements of the lower limbs using a handheld dynamometer and belt: a comparison with an isokinetic dynamometer. *J Phys Ther Sci*. 2011;23(4):553–557.
29. Matsudo SM, Matsudo VR, Araújo T, et al. Physical activity level of São Paulo State population: an analysis based on gender, age, socio-

economic status, demographics and knowledge. *Rev Bras Ciên Mov*. 2002;10(4):41–50.
30. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med*. 2002;21(9):1331–1335.
31. Kottner J, Audige L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48(6):661–671.
32. Fleiss JL. *The Design and Analysis of Clinical Experiments*. Hoboken, NJ: Wiley & Sons; 2003:7.
33. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005; 19(1):231–240.
34. Domholdt E. *Physical Therapy Research: Principles and Applications*. 2nd ed. Philadelphia, PA: WB Saunders; 2000.
35. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–310.
36. Bland M. *An Introduction to Medical Statistics*. 2nd ed. New York, NY: Oxford University Press; 1995.

Address correspondence to Débora Bevilaqua-Grossi, PhD, Ribeirão Preto Medical School, University of São Paulo, Avenida dos Bandeirantes, 3900–Monte Alegre, Ribeirão Preto, São Paulo, Brazil 14049-900. Address e-mail to deborabg@fmrp.usp.br.