Concurrent Validity of Lower Limb Muscle Strength by Handheld Dynamometry in Children 7 to 11 Years Old

Authors: Ryan Mahaffey; Megan Le Warne, Stewart C. Morrison, Wendy I. Drechsler, and Nichola Theis

Abstract

Context

The assessment of pediatric muscle strength is necessary in a range of applications, including rehabilitation programmes. Hand-held dynamometry is considered easy to use, portable and low cost, but validity to measure lower limb muscle strength in children has not been assessed.

Objective

To determine the concurrent validity of lower limb torque from hand-held dynamometry (HHD) compared to isokinetic dynamometry (ID) in children age 7 to 11 years old.

Design

A descriptive assessment of concurrent validity of lower limb joint torques from HHD compared to ID.

Methods

Sixty-one typically developing children underwent assessment of maximal hip, knee and ankle isometric torque by HHD and ID using standardized protocols. Joint positions were

selected to represent maximal strength and were replicated between devices. Concurrent validity was determined by Pearson's correlation, limits of agreement, and Bland-Altman plots.

Results

Correlations between HHD and ID were moderate-to-large for knee extension (*r* 95%CI: 0.39 to 0.73), small-to-large for plantarflexion (*r* 95%CI: 0.29 to 0.67), knee flexion (*r* 95%CI: 0.16 to 0.59), hip flexion (*r* 95%CI: 0.21 to 0.57), hip extension (*r* 95%CI: 0.18 to 0.54), and hip adduction (*r* 95%CI: 0.12 to 0.56), and small-to-moderate for dorsiflexion (*r* 95%CI: -0.11 to 0.39) and hip abduction (*r* 95%CI: -0.02 to 0.46). Limits of Agreement for all joint torques were greater than 10% indicating large error in HHD measured torque compared to ID. A positive proportional bias was detected for plantarflexion, indicating that HHD underestimated torque to a greater extent in participants with higher torque values.

Conclusions

Maximal torque values from HHD and ID are consistent with those previously reported in the literature. Poor concurrent validity of HHD may have arisen from issues around joint position, joint stabilization and the experience of the tester to prevent an isokinetic contraction. Pediatric lower limb muscle strength assessed by hand-held dynamometry should be interpreted with caution.

Introduction

Muscle strength is defined as the maximum tension a muscle or muscle group can exert during one voluntary action under specific conditions.[1] Muscle function, composed of strength, endurance and power, is an important fitness component for daily activities and functional tasks throughout the life span.[2] Typically, children exhibit a linear increase in muscle strength between the ages of 6 and 12 years old which is associated with gains in body size and improvements in movement skill aptitude.[2,3] A review by Smith et al.[3] reported inverse associations between paediatric muscle function (including muscle strength) and adiposity, cardiovascular disease and metabolic risk factors; and positive associations with bone health and self-esteem. Thus, the assessment of paediatric muscle strength is necessary in a range of applications including; defining the presence and severity of muscle weakness, examining the effectiveness of rehabilitation interventions and monitoring of training programmes.[1]

Two ways to objectively measure muscle strength are isokinetic dynamometers (ID) and hand-held dynamometers (HHD). Hand-held dynamometers are considered easy to use, portable, low cost and have been used in a range of paediatric studies to determine muscle strength.[4] Isokinetic dynamometer assessment has been considered the gold or reference

standard for measuring muscle strength due to its superior reliability compared to HHD.[1] Despite the expense and relative complexity, the use of IDs has become popular in sports, research and clinical settings, including testing of muscle strength in children.[1,5]

For a measurement tool to be deemed clinically useful, both validity and reliability need to be assessed in specific populations in which the tool is to be used. Criterion validity, i.e., how well an instrument measures what it intends to measure, can be assessed through concurrent collection of data from the instrument and a gold standard.[6] Few studies have assessed concurrent validity of HHD for use in children. Hebert et al. [7] assessed the concurrent validity of HHD compared to ID in adolescents, 13-17 years old. Intraclass correlation coefficients ranged from 0.48 to 0.93 across lower limb joints, with the lowest ICC's recorded for ankle plantarflexion and hip abduction, which may reflect the unstable position of measuring strength at these joints. While there are limitations to using only correlations to assess validity (e.g. systematic bias is not assessed),[8] these results generally showed that HHD was a valid measure of strength at the majority of joints. However, to the authors' knowledge, no study has examined the concurrent validity of HHD compared to ID in children younger than 13 years old, which is important to understand as 6-12 years represents a key stage of maturation to study strength changes in children.[2,3]

Therefore, the aim of this study was to determine the concurrent validity of maximal torque measures of lower limb joints from HHD compared to the gold standard ID in children age 7 to 11 years old.

Methods

Study Design

The study design was a descriptive assessment of concurrent validity, for which a sample size of >50 participants has been recommended.[5]

Participants

Sixty-six typically developing children (33 boys and 28 girls) from local primary schools were recruited to participate in the study. Schools gave permission for information and consent to be sent to parents/guardians for their children to participate. Mean and standard deviation (±) of the cohort were: age 9.20 ± 0.98 years, height 1.36 ± 0.07 m, weight 34.0. ± 8.0 kg, BMI z-score 0.51 ± 1.60. Parental/guardian informed written consent and child informed verbal assent was gathered prior to participation. Ethical approval was granted from the host institution. Inclusion criteria was constrained to participants willing and able to take part in strength assessment by HDD and ID. Exclusion criteria included any medical condition or injury affecting musculoskeletal, neuromuscular, or orthopaedic integrity.

Instrumentation and procedures

All testing took place in the institution's biomechanics laboratory in the presence of a teacher from the participants' school. Children wore shorts and t-shirt and removed shoes and socks during data collection. Hand-held (Wagner Force One FDIX Force Gauge, Wagner, Greenwich, CT USA) and isokinetic dynamometry (Cybex II; CSMI, Saughton, MA, USA) were performed to determine isometric strength of the hip abductors, adductors, flexors and extensors, knee flexors and extensors, and ankle dorsiflexors and plantarflexors. Both HHD and ID were measured in one session on the participant's dominant limb (determined by additional question on parental consent form). The order of testing between devices and

between lower limb joints was randomised for each participant. Participants' positions for HHD and ID are described in Table 1. The same protocols were implemented for ID and HHD (the "make" test was used for HHD).[1] Joint position on ID were selected to represent maximal strength joint position (approximately the middle of the range of motion).[9] Joint position during HHD were replicated to approximately match the same position during ID.

For both HHD and ID, stabilisation straps were applied tightly over the contralateral leg and torso and participants were instructed to cross their arms over their chest. An additional strap was applied to the waist for hip flexion and extension, and ankle plantarflexion and dorsiflexion. For knee flexion/extension a thigh strap was applied to both legs. Both the ID and HHD pads were placed in the same position on the limb (maximal distance from joint being tested immediately proximal to the distal joint). Verbal encouragement to push maximally was provided through the contraction. To familiarise the participants with the procedures, three submaximal isometric contractions were performed before each set of maximal contractions. A two-minute rest period was provided to minimise fatigue between warm-up and maximal efforts. Participants performed two 5-second maximal isometric contractions for each joint position, with 45-seconds rest in between trials. If peak force (HHD) or torque (ID) measured from the two maximal trials differed by 10% a third was performed and the mean of closest two trials was calculated.

Table 1.

Data analysis

Peak force, from HHD, was converted to torque by multiplying by the length of the lever arm (measured between the point of application of the dynamometer and the relevant joint centre). Joint centres were defined by bony landmarks of the greater trochanter, lateral femoral epicondyle, lateral malleolus. Isometric torque from Cybex ID was not corrected for limb weight to aid comparisons with clinical applications of HHD.[10] The HHD measurements were collected by one tester, with experience of collecting muscle strength data in children. The ID measurements were collected by a second tester, also with experience of collecting muscle strength data in children. Of the 66 participants, complete data set were recorded for 61. Due to time-constraints, two participants did not complete ID and three participants did not complete HHD. Only complete data sets from participants were analysed.

Statistical analysis

Statistical analyses were performed using SPSS (24.0; IBM Corp., Amonk, NY, USA). Data were assessed for normality (Shapiro-Wilk test), with all data conforming to a normal distribution. The validity of isometric torque measurement from HHD compared to ID was assessed in three ways. First, Pearson's correlation coefficients with 95% confidence intervals (95%CI) were performed. Correlation coefficients were interpreted in accordance with Cohen's magnitudes scale: $r < 0.1$ is trivial; $0.1 \leq r < 0.3$ is small; $0.3 \leq r < 0.5$ is moderate; $r \geq 0.5$ is large.[11] Second, agreement between measures was determined by Limits of Agreement (LoA), and Bland-Altman plots (differences in torque values between devices was determined to be normally distributed by Shapiro-Wilk test; see supplementary material). Correlations of the differences between HHD and ID torques and the average torque were examined to detect proportional bias ($r > 0.50$), which indicated the use of a

regression-based Bland-Altman plot. Third, to aid comparison with previous studies which validated HHD against ID, [7] ICC (3,1) and SEM were calculated for all isometric torques.

Results

Results for Pearson's correlation coefficients (95% CI), LoA, ICC and SEM are shown in Table 2. All LoA were greater than 10% indicating large error relative to the magnitude of torque recorded by ID compared to the HHD. Figure 1 presents the Bland-Altman plots for each joint tested.

Hip

Hip adduction, hip flexion and hip extension torque displayed small-to-large 95%CI correlations between HHD and ID. Hip abduction produced small-to-moderate 95%CI correlations between HHD and ID. No hip joint torques demonstrated a proportional bias between ID and HHD.

Knee

Knee extensor torque from HHD was the only variable to have moderate-to-large correlations with ID. Knee flexion displayed small-to-large 95%CI correlations between HHD and ID. No knee joint torques demonstrated a proportional bias between ID and HHD.

Ankle

Peak plantarflexion torque displayed small-to-large 95%CI correlations between HHD and ID. Peak dorsiflexion produced small-to-moderate 95%CI correlations between HHD and ID. A

positive proportional bias was detected for plantarflexion (r=0.84, slope 1.24, intercept -14.89, 95%CI 26.28 Nm), indicating that HHD underestimated torque to a greater extent in participants with higher torque values.

Table 2

Figure 1 of Bland Altman plots

Discussion

Concurrent validity of HHD, compared to ID, has been established in adults[12] and adolescents,[7] but not in children younger than 13 years old. The purpose of this study was to assess the validity of lower-limb torque measured from HHD compared to the gold standard ID in children 7-11 years old. Our results showed that knee extensor torque from HHD was the only variable to have moderate-to-large correlations with ID. Limits of agreement between HHD and ID ranged from 6.30 to 38.82 Nm or expressed as a percentage of torque measured from ID, ranged from 49.1 to 148.9%. All LoA were greater than 10% indicating large error relative to the magnitude of torque recorded by ID[5] demonstrating poor concurrent validity of HHD in the current study.

In comparison with validity of HHD in comparison to ID measured in adolescents,[7] the results from our study present lower ICC and SEM values across the same muscle groups (ICC 0.48 to 0.93 and SEM 0.5 to 6.0 Nm compared to ICC 0.25 to 0.71 and SEM 2.4 to 17.9 Nm in Herbert et al.[7] and the current study, respectively). Lower levels of validity found in

our younger cohort indicate that factors such as attention and understanding instructions are potential difficulties,[13] creating larger errors in torque measures from HHD. Low ICCs for ankle dorsiflexion and plantarflexion may reflect difficulties stabilising the ankle joint and maintaining 90 deg position.[7] Hip abduction ICCs were also low, which concurs from the findings of Herbert et al.[7] Indeed, Hebert et al.[7] recommended replicating their work on hip abduction validity due different positions being tested for HHD (supine) and ID (side lying). Our results confirm that differences in testing positioning does not account for low validity in this muscle groups (at least in younger children) and that perhaps the side-lying method itself, as an unstable position, may cause errors in muscle strength measurements.

Few studies have presented LoA for concurrent validation of HHD in comparison to ID. Mentiplay et al.[14] presented lower limb isometric torque from two HHD devices in comparison to ID in 30 adults. They reported lower LoA of for the hip (10-16 Nm) and knee (~20 Nm) compared to the current study (18-38 Nm and 24-31 Nm for hip and knee, respectively). In children aged 7-11 years, less aptitude to concentrate, stay on task, or follow instructions may explain the lower agreement between methods compared to an adult population.[7] However, LoA for the ankle were lower in the present study (6-24 Nm) compared to Mentiplay et al.[14] (~24 Nm). Furthermore, Mentiplay et al.[14] also found bias across plantarflexion, knee extension and hip flexion indicating that as torque increased, HHD underestimated torque to a greater extent. This may relate to the difficulty of the tester to stabilise and counter torque from larger musculature.[15] In the current study, bias was only present in plantarflexion, which may relate to issues stabilizing the ankle joint at 0 degrees throughout the test to match joint position on ID.[14]

To highlight the need to quantify concurrent validity when using HHD to determine normative, pathological and longitudinal strength changes in children, we compared the findings of our present study to previous literature. First, applying our findings to the sample of 7 to 11 years old (n= 62) from Eek et al.[16], shows that the LoA from the current study approximately overlaps the entire range of torque values from 7 to 11-year olds for plantarflexion, knee extension, knee flexion, and hip extension. Therefore, it is impossible to determine if a measurement of torque by HHD is considered average strength for their age or up to five years behind or in front compared to this normative database. Additionally, Hendengren et al.[4] reported HHD isometric joint torque in six children with juvenile chronic arthritis and six age- and gender-matched typically developing controls. Significant differences in torque were reported for dorsiflexion and plantarflexion. However, the mean difference between the patient and the control group was 4 Nm for both muscle groups. These small differences are within the LoA found in the current study (6.3 and 23.8 Nm for dorsiflexion and plantarflexion, respectively) indicating that HHD may not be sensitive enough to differentiate between patients and controls. Finally, an 8-week strength-based intervention in eight children with cerebral palsy (4-8 years old) utilising HHD to monitor changes in strength showed an increase in lower limb force of 26-88% from baseline to post intervention.[17] However, none of the increases across the lower limb joints were greater than the LoA% measured in the current study (49-149%). These comparisons with previous studies indicate the need for caution when using HHD assessment to determine normative values, compare across populations and monitor longitudinal changes. Further studies should be conducted to determine the validity and lower limb strength measurements using HHD by other assessors and in specific populations.

The concurrent validity between HHD and ID measures in previous studies has been assessed using several statistical approaches including correlation coefficients (Pearson's and ICCs), and LoA. In adult populations, correlations between torque measured from HHD and ID have ranged from 0.5-0.9,21 0.6-0.9,22 and as high as 0.99 for the knee extensors.[18] These values are considerably higher than the 0.2-0.6 reported in the current paediatric population but are likely the result of less heterogeneity in our sample compared to an adult population (e.g. Mentiplay et al.[14] reported standard deviations in their adult participants between 3.45 and 25.19 compared to between 2.67 and 18.98 in our paediatric cohort). Correlation based measures of validity are sensitive to sample heterogeneity (i.e. with a varied sample it is extremely easy to obtain a high value of r).[19] For this reason, correlation coefficients (including intra-class correlation coefficients) on their own are not sufficient for quantifying concurrent validity.[8] In addition, a correlation quantifies the degree to which two variables are related, but a high correlation does not automatically imply that there is good agreement between the two methods.[20] Thus, assessing validity through LoA is a more robust assessment of concurrent validity.

This study determined the concurrent validity of lower limb torque from HHD against the gold standard measure of ID in 7-11-year-old children. The results demonstrated poor validity for the majority of HHD torque measures and a general lack of agreement with ID in this age group. Issues around joint position, joint stabilisation and the experience of the tester to produce equal force to obtain an isometric rather than concentric/eccentric isokinetic contraction should all be considered when using HHD in paediatric populations.

The findings from our study adds to previous research on the validity of using HHD to measure muscle strength; showing that HHD are less valid in younger paediatric populations compared to previous research in adolescents and adults.

References

1. Jones M, & Stratton G. Muscle function assessment in children. *Acat Paediatr*. 2000;90:753-761

2. Beunen G, & Thomis M. Muscular Strength Development in Children and Adolescents. *Pediatr Exerc Sci*. 2000;12:174-197.

3. Smith JJ., Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. *Sports Med*. 2014;44(9):1209-1223

4. Hedengren E, Knutson L, Haglund-A Ê Kerlind Y, Hagelberg S. Lower extremity isometric joint torque in children with juvenile chronic arthritis. *Scand J Rheumatol*. 2001;30:69-76

5. Chamorro C, Armijo-Olivo S, De La Fuente C, Fuentes J, Javier Chirosa L. 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)*. 2017;12(1):359-375

6. van den Beld WA, van der Sanden GAC, Sengers RCA, Verbeek ALM., Gabreels FJM. Validity and reproducibility of hand-held dynamometry in children aged 4-11 years. *J Rehabil Med*. 2006;38:57-64

7. Hebert LJ, Maltais DB, Lepage C, Saulnier J, Crete M, Perron M. Isometric muscle strength in youth assessed by hand-held dynamometry: A feasibility, reliability, and validity study. *Pediatr Phys Ther*. 2011;23(3)289-299

8. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: A systematic review. *PLos ONE*. 2012;7(5)

9. Baltzopoulos P, King M, Gleeson N, De Croix, M. The BASES Expert Statement on Measurement of Muscle Strength with Isokinetic Dynamometry. Produced on behalf of the British Association of Sport and Exercise Sciences. Spring, 2012. https://www.bases.org.uk/imgs/baltz12_13_isokinetic49.pdf. Accessed February 2021.

10. Burnett CN, Filusch Betts E, King WM. Reliability of isokinetic measurements of hip muscle torque in young boys. *Physical Therapy*. 1990;70(4):244-249.

11. Cohen J. Statistical Power Analysis for the Behavioral Sciences (2nd ed.). New York: Routledge; 1988.

12. Stark T, Walker B, Philips 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.

13. Beenakker EAC, van der Hoevan JH, Fock JM, Maurits NM. Reference values of maximum isometric muscle force obtained in 270 children aged 4±16 years by hand-held dynamometry. *Neuromuscul Disord*. 2001;11: 441-446

14. Mentiplay BF, Parraton LG, Bower KJ, Adair B, Pua YH, Williams GP, McGae R, Clark RA. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: A reliability and validity study. *PLoS ONE*. 2015;10(10).

15. Wikholm J, & Bohannon R. Hand-held Dynamometer Measurements: Tester Strength. *J Orthop Sports Phys Ther*. 1991;13(4):191-198

16. Eek MN, Kroksmark AK, Beckung E. Isometric Muscle Torque in Children 5 to 15 Years of Age: Normative Data. *Arch Phys Med Rehabil*. 2006;87(8):1091-1099

17. Blundell SW, Shepherd RB, Dean CM, Adams RD, Cahill BM. Functional strength training in cerebral palsy: A pilot study of a group circuit training class for children aged 4-8 years. *Clin Rehabil*. 2003;17(1):48-57.

18. Piao C, Yoshimotoz N, Shitama H, Makino K, Wada F, Hachisuka K, Hachisuka K. Validity and Reliability of the Measurement of the Quardriceps Femoris Muscle Strength with a Hand-Held Dynamometer on the Affected Side in Hemiplegic Patients. *J UOEH*. 2004;26(1): 1-11

19. Atkinson G, & Nevill A. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998;26(4):217-238.

20. Giavarina D. Understanding Bland Altman analysis. *Biochemia Medica*. 2015;25(2)141-151.

Table 1. Summary of isometric testing muscle group, joint position angle and dynamometer set up position for ID and HHD. 0° represents full hip and knee extension, and neutral ankle in sagittal plane or neutral abduction/adduction in frontal plane.

Muscle group	Cybex ID	Wagner HHD
Ankle dorsiflexion	Supine	Prone
	Hip 90°	Hip 0°
	knee 90°	knee 0°
	ankle 0°	ankle 0°
Ankle	Supine	Prone
plantarflexion	Hip 90°	Hip 0°
	knee 90°	knee 0°
	ankle 0°	ankle 0°
Knee extension	Seated	Seated
	Hip 90°	Hip 90°
	Knee 60°	Knee 90°
Knee flexion	Seated	Seated
	Hip 90°	Hip 90°
	Knee30°	Knee 90°
Hip extension	Supine	Supine
	Hip 60°	Hip 60°
	Knee 90°	Knee 90°
Hip flexion	Supine	Supine
	Hip 30°	Hip 30°
	Knee 90°	Knee 90°
Hip abduction	Side lying	Side lying
	Hip 0°	Hip 0°
	Knee 0°	Knee 0°
Hip adduction	Side lying	Side lying
	Hip 20°	Hip 20°
	Knee 0°	Knee 0°

ID = Isokinetic Dynamometer

HHD = Hand-Held Dynamometer

Table 2. Mean ± Standard Deviation Isokinetic Dynamometer and Hand-Held Dynamometer Isometric Muscle Torque, and Validity Statistics

Muscle group	HHD Mean ± SD (Nm)	ID Mean ± SD (Nm)	r (95%CI)	LoA (Nm)	LoA% (of ID) (Nm)	Proportional bias (r)	ICC 3,1 (95% CI)	SEM (Nm)
Ankle dorsiflexion	9.23 ± 2.23	6.83 ± 2.67	0.15 (-0.11 to 0.39)	6.3	92.26	0.18	0.25 (-0.24 to 0.55)	2.36
Ankle plantarflexion	17.57 ± 4.53	35.64 ± 13.80	0.51 (0.29 to 0.67)	23.81	66.81	0.84*	0.46 (0.10 to 0.68)	10.03
Knee extension	57.58 ± 13.29	62.25 ± 18.98	0.58 (0.39 to 0.73)	30.57	49.11	0.39	0.71 (0.51 to 0.82)	8.92
Knee flexion	37.48 ± 12.42	33.32 ± 9.83	0.39 (0.16 to 0.59)	24.37	73.12	-0.26	0.56 (0.26 to 0.73)	7.57
Hip extension	67.72 ± 19.71	26.07 ± 15.06	0.38 (0.18 to 0.54)	38.82	148.89	0.28	0.57 (0.15 to 0.74)	17.87
Hip flexion	33.69 ± 6.74	19.43 ± 9.81	0.40 (0.21 to 0.57)	18.45	94.94	0.39	0.64 (0.30 to 0.79)	6.61
Hip abduction	39.48 ± 9.41	15.98 ± 6.66	0.23 (-0.02 to 0.46)	19.97	124.96	0.34	0.36 (-0.07 to 0.62)	11.46
Hip adduction	43.75 ± 11.40	20.99 ± 10.43	0.36 (0.12 to 0.56)	24.19	115.25	-0.10	0.53 (0.22 to 0.72)	10.80

* denotes proportional bias detected (r > 0.5)

ID = Isokinetic Dynamometer

HHD = Hand-Held Dynamometer

LoA = Limits of Agreement

CI = Confidence Interval

r = Pearson’s Correlation Coefficient

ICC = Intraclass correlation coefficient

SEM = Standard Error of Measurement

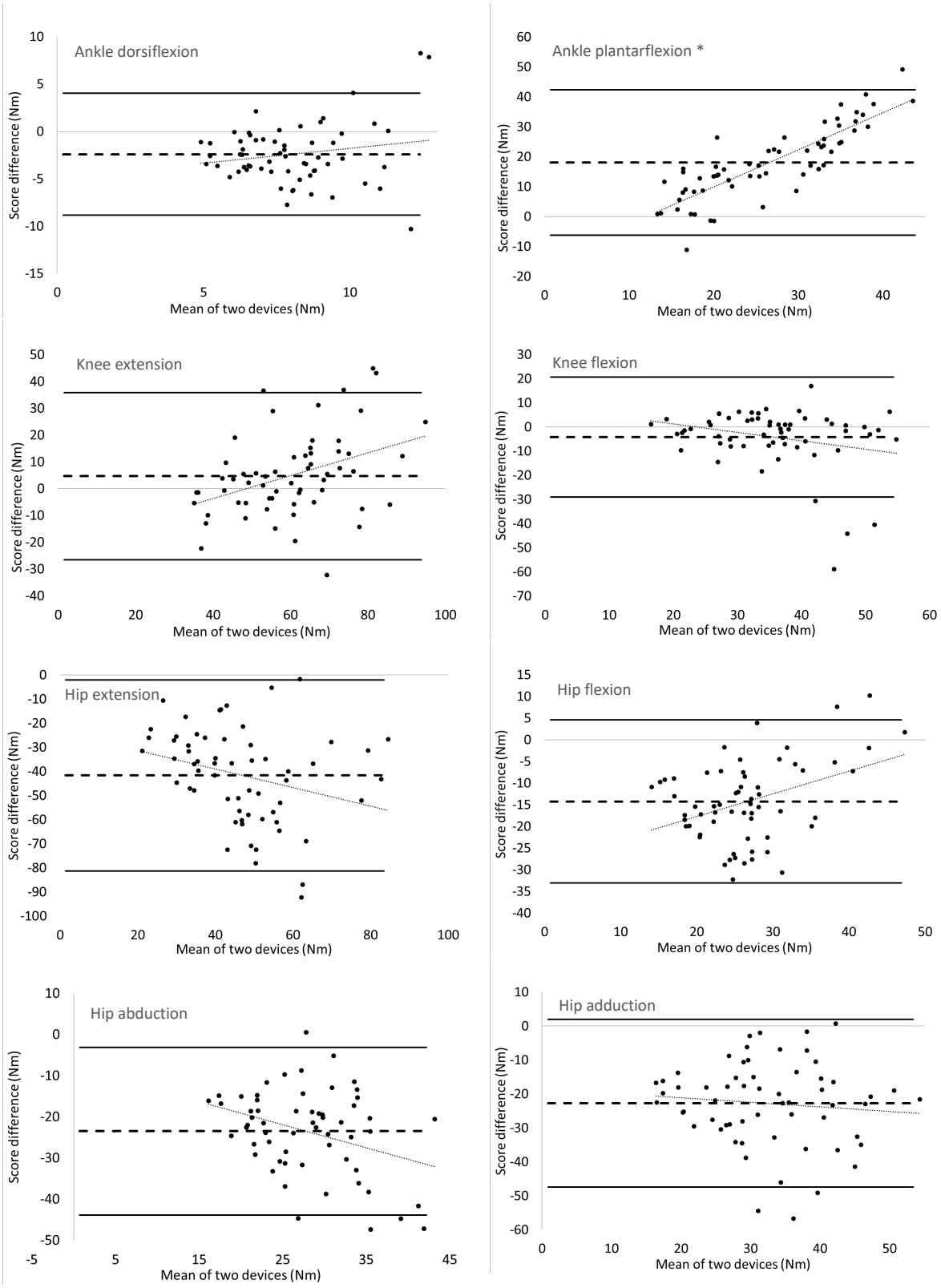
Ankle dorsiflexion
Score difference (Nm)
Mean of two devices (Nm)
Ankle plantarflexion *
Score difference (Nm)
Mean of two devices (Nm)
Knee extension
Score difference (Nm)
Mean of two devices (Nm)
Knee flexion
Score difference (Nm)
Mean of two devices (Nm)
Hip extension
Score difference (Nm)
Mean of two devices (Nm)
Hip flexion
Score difference (Nm)
Mean of two devices (Nm)
Hip abduction
Score difference (Nm)
Mean of two devices (Nm)
Hip adduction
Score difference (Nm)
Mean of two devices (Nm)

Figure 1. Bland-Altman Plots of Isokinetic dynamometer and Hand-Held Dynamometer Lower Limb Torques. Dashed horizontal line represents mean (score) difference between devices. Black horizontal lines represent 95% confidence interval of the mean (score) differences between devices. Dotted line represents correlation between mean difference of devices and mean torque obtained from devices (proportional bias defined by $r > 0.5$, denoted by *).