

Research article

Claudio Chamorro*, Susan Armijo-Olivo, Carlos De la Fuente, Javiera Fuentes, Luis Javier Chiroso

Absolute reliability and concurrent validity of hand held dynamometry and isokinetic dynamometry in the hip, knee and ankle joint: systematic review and meta-analysis

<https://doi.org/10.1515/med-2017-0052>

received April 21, 2017; accepted August 14, 2017

Abstract: The purpose of the study is to establish absolute reliability and concurrent validity between hand-held dynamometers (HHDs) and isokinetic dynamometers (IDs) in lower extremity peak torque assessment. Medline, Embase, CINAHL databases were searched for studies related to psychometric properties in muscle dynamometry. Studies considering standard error of measurement SEM (%) or limit of agreement LOA (%) expressed as percentage of the mean, were considered to establish absolute reliability while studies using intra-class correlation coefficient (ICC) were considered to establish concurrent validity between dynamometers. In total, 17 studies were included in the meta-analysis. The COSMIN checklist classified them between fair and poor. Using HHDs, knee extension LOA (%) was 33.59%, 95% confidence interval (CI) 23.91 to 43.26 and ankle plantar flexion LOA (%) was 48.87%, CI 35.19 to 62.56. Using IDs, hip adduction and extension; knee flexion and extension; and ankle dorsi-

flexion showed LOA (%) under 15%. Lower hip, knee, and ankle LOA (%) were obtained using an ID compared to HHD. ICC between devices ranged between 0.62, CI (0.37 to 0.87) for ankle dorsiflexion to 0.94, IC (0.91 to 0.98) for hip adduction. Very high correlation were found for hip adductors and hip flexors and moderate correlations for knee flexors/extensors and ankle plantar/dorsiflexors.

Keywords: Lower extremities; Muscle strength; Reproducibility of results

1 Introduction

Assessing muscle strength is an important clinical consideration for patients who may have a neurological, muscular, and/or skeletal illness [1,2]. Muscle force assessments are commonly performed before and after interventions to quantify treatment effectiveness [3]. The psychometric properties of strength devices are important not only for research but also for clinical practice. The ability to determine if a device is valid, reliable, and/or responsive in a determined context can help clinicians decide when and how to use it. Two ways to objectively measure muscle strength are isokinetic dynamometers (IDs) and hand-held dynamometers (HHDs). While the psychometric properties of these devices have been investigated in different contexts using different models, joints and conditions, the resulting information is fragmented and difficult to comprehensively understand [4-8].

HHDs provide a quantified measurement of force. They are considered easy use, with convenient size, and low cost. The overall affordability of this device may justify further widespread clinical use but reported reliability of HHDs for measuring lower-extremity strength differs widely between authors. For example, Kelln et al. [9]

*Corresponding author: Claudio Chamorro, Carrera de Kinesiología, Escuela de Medicina, Edificio Ciencias de la Salud, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile. Phone number 56223541326, E-mail: cchamorro@uc.cl.

Claudio Chamorro, Servicio Kinesiología, Clínica UC San Carlos de Apoquindo, Santiago, Chile

Susan Armijo-Olivo, Faculty of Rehabilitation Science, University of Alberta, Edmonton Canada

Carlos De la Fuente, Javiera Fuentes, Carrera de Kinesiología, UDA, Cs de la Salud, Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile

Carlos De la Fuente, Facultad Cs. de la Rehabilitación, Universidad Andrés Bello, Santiago, Chile

Luis Javier Chiroso, Departamento de Educación Física y Deportiva, Universidad de Granada, Granada, España

reports a standard error of measurement expressed as percentage of the mean (SEM%) of 4% in knee flexors strength assessment while Lu *et al.* [10] reports SEM% of 14%. Similarly, in relation to hip abductors strength assessment where Kelln *et al.* [9] reports SEM% of only 1% while Arnold *et al.* [11] reports 21%.

The use of IDs has become progressively popular in sports, research, and clinical settings [12]. The reliable test results, particularly for muscles of the lower extremity, have made IDs the gold standard for measuring muscle strength [13] mainly because the results are not influenced by a strength imbalance between the participant and the examiner, whereby a maximal torque can be generated throughout the whole range of motion [14]. Indeed, IDs provide mechanically valid and reliable measures of torque, position, and velocity for both clinical and research purposes [15]. Nevertheless, the elevated costs of this device limit widespread use in clinical practice. Although considered the gold standard, differences in SEM (%) also exists between authors in lower extremity strength assessment. For example Holmack *et al.* [16] reports twice the value of SEM (%) in comparison to Morrison *et al.* [17] reports in ankle dorsiflexion strength assessment.

To the best of our knowledge, no systematic review currently exists that has summarized the psychometric properties of these devices, particularly in regards to absolute reliability and concurrent validity. The purposes of this systematic review were to 1) examine absolute reliability using the standard error of measurement (SEM) and limit of agreement (LOA) in the hip, knee and ankle joint and 2) determine the concurrent validity between the HHDs and IDs in the joints just mentioned.

2 Methods

The reporting of this systematic review is based on Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines [18]. The PRISMA guidelines consist of a 27-item checklist and 4-phase flow diagram.

2.1 Search strategy

A search was performed for relevant studies from 1987 up to and including November 2016 related to the psychometric properties of muscle dynamometry. For this, several bibliographic databases were extensively explored,

including Medline, Embase, CINAHL, and the ISI Web of Science. The following words and combinations thereof were included as search terms: dynamometry, muscle strength, power or torque, isokinetic device, machine or instrument, reliability, validity, inter class or interclass, intra class or intra-class, inter rater or intra rater, inter tester, intra examiner, sensitivity or specificity, and gold standard. This database search was complemented by manually checking the bibliographies of identified papers for relevant key authors and journals. The search strategy was guided by a trained librarian.

2.2 Study selection

The inclusion criteria for studies assessed in this review were 1) included asymptomatic participants; 2) evaluated the dominant side of participants using isometric or concentric contractions (60°/s angular velocity) with either a HHD or ID in any of the joints of interest (i.e. hip, knee, or ankle); and 3) considered the following psychometric properties for the HHD and ID: a) Absolute reliability, expressed as the SEM and LOA for within subject variability between trials, and b) Concurrent validity, expressed as the inter device correlation coefficient (ICC) and 95% confidence interval (CI).

The exclusion criteria for this review were 1) studies published in a language other than English; 2) studies that analyzed only relative reliability (i.e. ICC) but not absolute reliability; 3) concurrent validity was expressed as Pearson correlation instead of ICC.

2.2.1 Definition of psychometric properties

The following definitions for reliability and concurrent validity were used.

2.2.1.1 Absolute reliability

Absolute reliability is the degree to which repeated measurements vary for individuals. The less repeated measurements vary, the higher the reliability. Absolute reliability is expressed either in the actual units of measurement or as a proportion of the measured values (i.e. dimensionless ratio). The most common method for analyzing absolute reliability is the SEM or LOA for within-subject variation [19,20]. The SEM quantifies score reliability within individual participants on different occasions. To produce a unit-free indicator of SEM error magnitude, the results can be expressed as SEM%. LOA, in turn, provides a value

range within which a truly unchanged participant score would be expected to remain over repeat testing, at a 95%CI [19]. To produce a unit-free indicator, LOA can also be expressed in percentage (LOA%)

2.2.1.2 Concurrent validity

This property measures how well a new instrument compares to a well-established gold standard [21]. The most common method for analyzing concurrent validity is the ICC. The ICC reflects a test's ability to differentiate between participants and, hence, the position of the individual relative to others in the group. However, the ICC does not provide information about the accuracy of individual scores [19].

2.3 Data extraction and quality assessment

Two independent reviewers (CCH and JF) screened the abstracts/titles of the publications found in the databases. After initial selection, the reviewers then further analyzed each paper based on inclusion/exclusion criteria. Each criterion was graded on a yes/no basis. If discrepancies existed between reviewers regarding a particular paper meeting a criterion, the ratings were compared and discussed until a consensus was reached.

The following characteristics were extracted from studies analyzing the absolute reliability of hand-held and isokinetic dynamometry: i) author, ii) publication year, iii) sample characteristics, iv) type of dynamometer used, v) joint and movement studied, vi) assessment position, vii) type of muscle contraction, viii) SEM and LOA expressed as absolute values and as percentages of the mean value between peak muscle assessments (i.e. tests 1 and 2). The following characteristics were extracted from studies analyzing concurrent validity between HHDs and IDs: i) author, ii) publication year, iii) sample characteristics, iv) type of dynamometer used, v) ID used as gold standard, vi) joint and movement studied, vii) assessment position and viii) $ICC_{2,1}$ based on a 2-way random effects repeated measures analysis of variance model.

3 Quality assessment methodology

The COSMIN checklist stands for Consensus-based standards for the selection of health measurements instruments. It is a recognized valid tool for evaluating the psychometric properties of health instruments [22]. Form C,

for absolute reliability studies, and form G, for concurrent validity studies, were used.

The COSMIN checklist for the absolute reliability included 11 items. These items assessed if the sample size was appropriate, if there was a description of missing values and how they were handled, if there was an independent administration of each test, and if the time interval between tests was stated and appropriate, among others. In the case of concurrent validity, six items were considered in the COSMIN checklist, sample size, description and handling of missing data, and if the used criterion could be considered a gold standard, randomization process and independency in measurements.

Each item for each form was assigned a score of excellent, good, fair, or poor. In cases where the study did not consider a particular item, the item was listed as non-applicable. The methodological quality per study and form was obtained by considering the lowest rated item. For example, if one item on the "reliability" form was rated "poor," the methodological quality of that reliability study was also rated "poor."

All critical appraisals were independently completed by the two reviewers, and the results were compared. Any discrepancies were settled through discussion.

3.1 Data synthesis and analysis

Studies investigating similar outcomes (i.e. LOA, ICC) and those providing clear quantitative data were grouped, evaluated for heterogeneity, and pooled if possible. A meta-analysis was performed to quantify the pooled absolute reliability (i.e. LOA for within subject variation) of hand-held and isokinetic dynamometry in assessing muscle force in the hip, knee, and ankle joints.

Concurrent validity was quantified and expressed as the inter-machine ICC for the HHD as compared to isokinetic testing, the gold standard for assessing muscle strength. ICC was based on a 2-way random effects repeated measures analysis of variance model with absolute agreement. Munro's scale was selected to determine the level of agreement between devices, where 0.0-0.25 represented little correlation; 0.26-0.49 low correlation; 0.50-0.69 moderate correlation; 0.70-0.89 high correlation; and 0.90-1.0 very high correlation [23].

The Stata 13 software was used to pool effects and construct the forest plots for all comparisons. This analysis used a 95% CI. A test for heterogeneity was performed using a Chi-square test ($p < 0.10$). If clinical heterogeneity existed in the study population or intervention, the DerSimonian and Laird Random Effects Model of Pooling was

used based on the assumption of inter-study variability, thus providing more conservative estimates of the true effect. The SEM results were not pooled since, as a measure of variability, 95%CI cannot be computed. However, SEM data were described in terms of minimal, maximum, and range of numbers between studies.

4 Results

A total of 7920 articles were initially found through the database search. Of these, 138 were selected as potential studies of interest based on the abstract and title review (Figure 1). After full article screening, only 30 studies met the selection criteria. The kappa agreement between the reviewers in selecting articles after applying the inclusion and exclusion criteria was $k = 0.9$.

4.1 Characteristics of the studies

4.1.1 General results

Of the 30 studies that met initial selection criteria, 17 provided enough data for inclusion in the meta-analysis. The remaining 13 were not included either due to a lack of CI for the concurrent validity ICC or due to values not being shown in kg for HHD assessments or in Nm for ID assessments. Fifteen of the assessed studies [9,11,16,17,24-34] provided a detailed review for the LOA (%) of within-subject variations between trials 1 and 2 for the HHD and ID while assessing muscle force in the hip, knee, and ankle joints. Characteristics and outcomes of selected studies included in meta-analysis analyzing absolute reliability of HHD and ID are shown in Table 1 and 2. Characteristics and outcomes of selected studies included in meta-analysis analyzing concurrent validity between HHD and ID are shown in Table 3 and 4.

4.2 Methodological quality of the studies

4.2.1 Absolute reliability

The results of the critical appraisal for the selected studies that analyzed absolute reliability are presented in Table 5. Three [16,24,32] of the fifteen studies presented fair methodological quality, and twelve [9,11,17,25-31,33,34] presented poor quality. According to the COSMIN quality assessment [22], methodological quality is obtained by

considering the lowest rating of any item (i.e. worst score count). The poor methodological quality of the studies was mainly due to a low sample size (< 20). Indeed, only four [16,24,26,32] studies scored fair on this point. Another important methodological flaw was the assessed time interval. To maintain the independence of administrations, there should normally be at least seven days between measurements [37]. In the assessed studies, the time interval ranged from 1 minute to 14 days; with three studies classified as fair [9,11,30] and two as poor [25,26]. Although receiving an overall poor score, five studies [9,11,25,26,30] had either good or excellent methodological quality regarding the randomization process. Randomization was not applicable in three studies [16,32,34]. The percentage of missing values was not given in any study, and test conditions were similar in all studies.

4.2.2 Concurrent validity

The results of the critical appraisal for the selected studies that analyzed concurrent validity are presented in Table 6. Two [11,36] were scored as poor because of a low sample size (< 20). Another two [26,35] were scored as fair regarding methodological quality. No study mentioned how missing values were handled. Generally, the ID used in each study was explicitly mentioned and recognized as the gold standard; except for one [11] study in which it is left assumed, but not mentioned, that isokinetic dynamometry is the gold standard.

4.3 Meta-analysis results

Seventeen studies were included in the meta-analysis (Figure 1). For HHDs, absolute reliability assessments, kilograms (kg) was considered as unit of measure while Newton*meter (Nm) was considered for IDs analysis. Concurrent validity was expressed as the ICC in all selected studies. Only studies that presented HHD peak torque in Nm were considered, thereby facilitating comparison to ID studies.

4.3.1 HHD: absolute reliability

Results are shown in Table 7 represents sizes expressed in kg (%). One study [26] simultaneously assessed absolute reliability of the Lafayette and Hoggan Health HHDs in the hip, knee, and ankle.

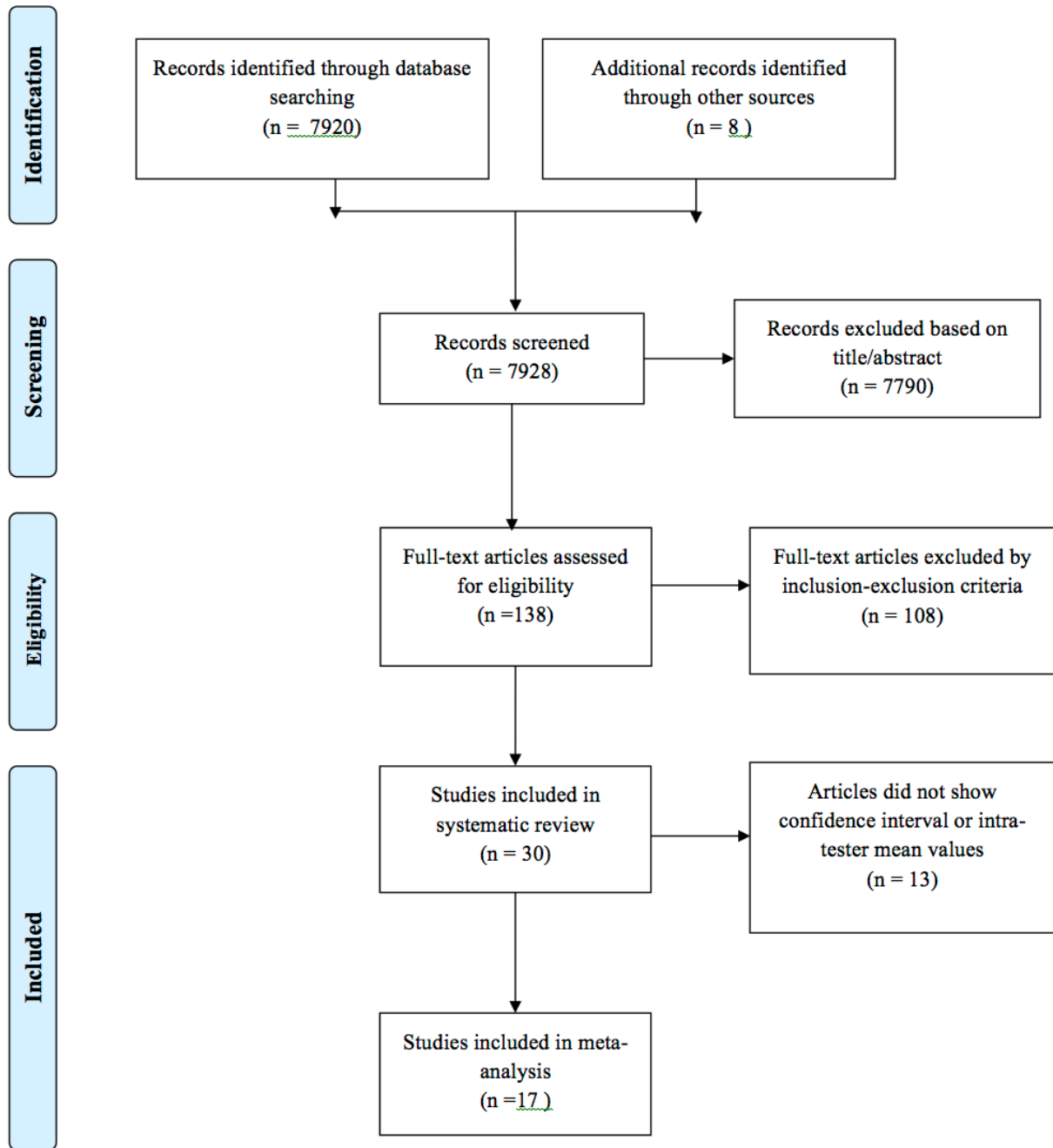


Figure 1: Flow-chart representation of selection process for manuscripts to be considered within this systematic review

Hip abduction

Four studies were included for hip abduction analysis [9,11,25,26]. Three of the studies used the Hoggan Health HHD [9,25,26] and two used the Lafayette HHD. [11,26] Two did assessments in a standing position [11,25], while two studies used the supine position [9,26].

Hip adduction

Two studies were included hip adduction analysis [9,26]. One used a Hoggan Health HHD [9] while the other used both the Hoggan Health and Lafayette HHDs [26]. All assessments were done in a supine position.

Table 1: Characteristics of Selected Studies Analyzing Absolute Reliability of HHD and ID.

Author	Sample	Type of dynamometer	Joint and movement	Assessment position	Type of muscle contraction
Arnold, C	18 YHA	Lafayette, HHD	Hip: flex, abd, and ext Knee: ext Ankle: DF	Hip flex, abd: Standing Hip ext: Prone Knee: Seated Ankle: Seated	Isometric
Kelln, B	20 HA	Microfet 2, Hoggan Health Industries, HHD	Hip: add, abd, flex, ext, Knee: flex, ext Ankle: DF	Hip and Knee ext: Prone Hip flex, add, abd and Ankle DF: Supine	Isometric
Kawaguchi, J	8 YHA	Fet 2 Hoggan Health Industries, HHD	Hip: abd, ext	Standing	Isometric
Mentiplay, B	30 HA	Lafayette, Hoggan Health Industries, HHD; Kin Com, ID	Ankle: DF, PF Knee: flex, ext Hip: abd-add; flex, ext	Hip flex and knee ext, flex: Seated Ankle DF, PF and Hip add, abd: Supine Hip ext: Prone:	HHD: isometric ID: concentric 60°/s
Tung Wu, L	25 HA	GT-10, HHD	Knee: flex, ext	Seated	Isometric
Claiborne, T	13 HA	Biodex, ID	Hip: add, abd, flex, ext,	Hip add/abd: Supine Hip flex, ext: Standing	Concentric 60°/s
Dauty, M	10 YHA	Cybex, 6000, ID	Knee: flexion	Seated	Concentric 60°/s
Harmann, A	24 adults over 70	Biodex; ID	Knee: flex, ext Ankle: DF, PF	Knee flex, ext: Seated Ankle DF, PF: Supine	Concentric 60°/s
Pereira de Carvalho, A	20 HA	Rev 9000; ID	Knee: flex, ext	Seated	Isometric
Ferri Morales, A	40 HA	Biodex, ID	Knee: ext	Seated	Isometric Concentric 60°/s
Dervisevic, C	16 adult men	Rev 9000, ID	Knee: flex, ext	Seated	Concentric 60°/s
Larsson, B	20 men	Kin Com, ID	Knee: ext	Seated	Concentric 60°/s
Morrison, K	26 YHA	Kin Com, ID	Ankle: DF, PF	Seated	Concentric 60°/s
Holmback, A	15 HA	Biodex, ID	Ankle: DF	Seated	Concentric 60°/s
Kim, WK	27 female	J Tech, HHD	Knee: ext	Seated	Isometric

Abbreviations: abd, abduction; add, adduction; DF, dorsiflexion; ext, extension; flex, flexion; HA, healthy adults; HHD, hand-held dynamometry; ID, isokinetic dynamometry; PF, plantar flexion; YHA, young healthy adults.

Hip flexion

Three studies were included for hip flexion analysis. [9,11,26] One used a Hoggan Health HHD [9]; one used a Lafayette HHD [11]; and one used both the Hoggan Health and Lafayette HHDs [26]; two studies assessed hip flexion while seated, [9,26] and the third assessed flexion while standing. [11]

Hip extension

Four studies were included for hip extension analysis. [9,11,25,26]. Three studies used the Hoggan Health HHD [9,25,26] and two used the Lafayette HHD [11,26]. All assessments were done in a prone position.

Knee flexion

Three studies were included for knee flexion analysis [9,26,27]. One study used the Hoggan Health HHD [9]

Table 2: Outcome studies analyzing absolute reliability of hhd and id.

Author	N	Joint/Dynamometer	Movement (Strength Unit)	Intra-rater Mean Force Assessment	SEM (SEM%)	Difference Scores between Trials (%)	LOA Lower Limit (%)	LOA Upper Limit (%)	
Arnold, C	18	Hip / HHD	Flex: (kg)	8.59	1.81 (21)	6.5	-51.8	64.8	
			Abd: (kg)	9.89	1.07 (10.8)	2	-27.9	31.99	
			Ext: (kg)	10.43	1.89 (18)	1	-49.2	51.15	
	18	Knee / HHD	Ext (kg)	11.3	1.67 (14.7)	11.9	-28.8	52.71	
			Ankle / HHD	DF (kg)	5.6	1.36 (14.8)	5.6	-35.4	46.81
Kelln, B	20	Hip / HHD	Ankle / HHD	DF(kg)	9.2	0.45 (2.2)	9.3	3.2	15.38
			Hip add (kg)	10.35	0.24 (2.4)	4.8	-1.6	11.2	
			Hip abd (kg)	12.45	0.22 (1.77)	8.8	3.93	13.74	
			Hip flex (kg)	13.2	0.1 (0.83)	1.5	-0.78	3.81	
		Knee / HHD	Hip ext (kg)	13.25	0.64 (4.8)	15.8	2.47	29.23	
			Knee flex (kg)	12.95	0.48 (1.93)	2.3	-8.07	12.7	
			Knee ext (kg)	26	1.0 (4.0)	16.2	4.94	27.3	
Kawaguchi, J	8	Hip / HHD	Ext (kg)	29.65	1.57 (6.76)	1.9	-28	36.8	
			Abd (kg)	36.5	5.1 (14.1)	6.3	-32.7	45.4	
Mentiplay, B	30	Ankle / HHD	DF, Lafayette (kg)	18.5	1.4 (8.59)	7.3	-16.5	31.22	
			DF, Hoggan (kg)	20.9	1.29 (6.2)	0.14	-17.0	17.32	
			PF, Lafayette (kg)	50.71	4.3 (8.52)	1.1	-22.4	24.77	
			PF, Hoggan (kg)	47.9	3.0 (6.0)	0.5	-16.3	17.27	
		Hip / HHD	Hip abd, Lafayette (kg)	13.4	1.3 (9.7)	6.2	-20.3	32.82	
			Hip abd, Hoggan (kg)	13.0	1.4 (7.2)	3.5	-16.6	23.76	
		Hip / ID	Hip abd, Kin Com (Nm)	57.35	3.5(6.17)	7.8	-9.2	24.9	
			Hip/ HHD	Hip add, Lafayette (kg)	18.57	1.25 (6.8)	1.6	-17.2	19.43
		Hip add, Hoggan (kg)		18.16	1.02 (5.7)	3.5	-12.3	19.43	
		Hip / ID	Hip add, Kin Com (Nm)	94.7	3.4 (4.48)	0.52	-11.9	12.9	
			Hip/ HHD	Hip ext, Lafayette (kg)	23.23	1.5 (5.19)	0.8	-16.8	20.65
		Hip ext, Hoggan (kg)		23.5	1.2 (11.7)	4.4	-28.0	36.79	
		Hip/ ID	Hip ext, Kin Com (Nm)	128	8.99 (7.03)	1.5	-17.9	21.03	
			Hip/ HHD	Hip flex, Lafayett (kg)	30.2	1.8 (6.15)	1.3	-15.7	18.32
Hip flex, Hoggan (kg)	31.4	1.7 (5.3)		1.6	-13.4	16.6			
Hip / ID	Hip flex, Kin Com (Nm)	93.75	6.05 (6.45)	6.93	-10.9	24.8			

continued Table 2: Outcome studies analyzing absolute reliability of hhd and id.

			Knee ext, Lafayette (kg)	42.8	3.4 (7.72)	6.4	-14.9	27.85
		Knee / HHD	Knee ext, Hoggan (kg)	48.2	4.25 (8.52)	8.9	-14.6	32.65
		Knee / ID	Knee ext, Kin Com (Nm)	305.2	17.3 (5.67)	8.02	-7.67	23.7
		Knee / HHD	Knee flex, Lafayette (kg)	23.73	1.6 (6.93)	0.38	-18.8	19.58
			Knee flex, Hoggan (kg)	23.78	2.04 (8.58)	1.7	-22.0	25.52
		Knee / ID	Knee flex, Kin Com (Nm)	128.7	8.58 (6.67)	0.4	-18.1	18.8
Kim, WK	27	Knee / HHD	Ext (Nm)	46.76	2.28 (4.87)	1.8	-12.0	15.28
			Abd (Nm)	123.6	13.5 (10.9)	4.85	-25.5	35.2
		Hip / ID	Add (Nm)	128.5	24.1 (18.7)	9.0	-42.9	60.9
			Flex (Nm)	122	13.5 (10.7)	18.4	-11.3	48.2
			Ext (Nm)	140	10.4 (7.4)	0.5	-20.0	21.15
Dauty, M	10	Knee / ID	Flex (Nm)	124.35	7.0 (5.63)	1.36	-14.2	16.9
			Flex (Nm)	41.35	5.3 (12.8)	4.6	-30.8	40.1
		Knee / ID	Ext (Nm)	88.8	7.1 (8.0)	0.22	-21.9	22.3
Harmann, A	24		DF (Nm)	10.3	0.9 (8.74)	1.9	-22.3	26.2
		Ankle / ID	PF (Nm)	41.8	5.7 (13.8)	3.8	-34.6	42.3
			Ext (Nm)	226	8.1 (3.6)	2.6	-7.32	12.6
Pereira de Carvalho, A	20	Knee / ID	Flex (Nm)	127	6.2 (4.9)	3.1	-10.4	16.7
Ferri Morales, A	40	Knee / ID	Knee (Nm)	124.9	4.8 (2.48)	1.9	-8.9	12.7
			Ext (Nm)	142.5	15.5 (10.9)	4.9	-25.3	35.1
Dervisevic, E	16	Knee / ID	Flex (Nm)	97.9	15.5 (15.8)	2.5	-41.4	46.5
Larson, B	20	Knee / ID	Ext (Nm)	182.5	13 (7.12)	4.9	-14.8	24.6
			PF(Nm)	150.2	14.7 (9.8)	0.7	-26.4	27.8
Morrison, K	26	Ankle / ID	DF(Nm)	41.9	4.58 (10.9)	10.3	-19.8	40.6
Holmback, A	30	Ankle / ID	DF (Nm)	23.7	1.4 (5.9)	0.84	-15.5	17.2
			Flex (Kg)	27.55	3.72 (13.5)	1.8	-35.5	39.21
Tung Wu, Lu	25	Knee / HHD	Ext (kg)	36.35	2.2 (6.1)	0.8	-16.3	17.9

Abbreviations: abd, abduction; add, adduction; DF, dorsiflexion; ER, external rotation; ext, extension; flex, flexion; HHD, hand-held dynamometry; ID, isokinetic dynamometry; PF, plantar flexion.

Table 3: Characteristics of selected studies analyzing concurrent validity between HHD and ID.

Author	Sample	Joint and movement	Assessment position	Dynamometer	Compared to	Type of muscle contraction
Arnold, C	18 YHA	Hip: flex, abd, ext Knee: ext Ankle: DF	Hip flex, abd: Standing Knee ext and Ankle DF: Seated Hip ext: Prone	Lafayette, HHD	Biodex, ID	Isometric
Mentiplay, B	30 HA	Ankle: DF, PF Knee: flex, ext Hip: abd, add, flex, ext	Hip flex and Knee flex, ext: Seated Ankle PF, DF and Hip add, abd: Supine Hip ext: Prone	Lafayette, Hoggan HHD	Kin Com, ID	Isometric
Neil, S	10 YHA	Knee: flex, ext	Seated	HUR and performance recorder (PR 1) with external resistance	Biodex, ID	Isometric
Wang, Y	68 HA	Knee: ext	Seated	Integrated load cell device with and without external fixation	Biodex, ID	Isometric

Abbreviations: abd, abduction; add, adduction; DF, dorsiflexion; ext, extension; flex, flexion; HA, healthy adults; HHD, hand-held dynamometry; ID, isokinetic dynamometry; PF, plantar flexion; YHA, young healthy adults.

while the other two studies used the Lafayette HHD [26] or the GT-10 HHD. [27] Two studies carried out assessments in seated position [26,27] and the third study used a prone position. [9]

Knee extension

Five studies were included for knee extension analysis. [9,11,24,26,27] The GT-10 [27], J Tech [24], Lafayette [11,26] and Hoggan Health HHDs [9,26] were used. Four studies used a seated position [11,24,26,27] and one used a prone position. [9]

Ankle plantar flexion

A single study was found assessing ankle plantar flexion, but this study assessed both the Hoggan Health and Lafayette HHDs in a supine position [26].

Ankle dorsiflexion

Three studies were included for ankle dorsiflexion analysis. [9,11,26] One included the Hoggan Health HHD, [9] one

the Lafayette HHD [11] and one assessed both HHDs. [26] Two studies analyzed ankle dorsiflexion in a supine position [9,26] and the third study used a seated position. [11]

4.3.2 ID: absolute reliability (Table 8)

The results for absolute reliability comparisons (Nm %) using isokinetic dynamometry in the hip, knee, and ankle joints are displayed in Table 8.

Hip abduction and adduction. Two studies were included for hip abduction and adduction analyses [26,28]. One [26] assessed the movements in a supine position using the Kin Com ID, whereas the second study used a supine position with the Biodex ID [28].

Hip flexion and extension. Two studies were included for hip flexion and extension analyses [26,28]. One study assessed these movements in a standing position using the Biodex ID [28]. The second study assessed hip flexion in a sitting position and hip extension in a prone position, both using the Kin Com device [26].

Knee flexion and extension. Five studies were included for knee flexion analysis [26,29-31,33]. The Rev 9000 [31,33], Cybex [29], Kin Com [26] and Biodex [30] devices were used. Six studies were included for knee

Table 4: Outcomes of selected studies analyzing concurrent validity between HHD and ID.

Author	N	Analysis	Intermachine ICC
Arnold, C	18	Hip flex	0.70 (0.06-0.8)
		Knee ext	0.44 (0.05-0.76)
		Hip ext	0.52 (0.09-0.79)
		Hip ab	0.7 (0.34-0.88)
Wang, Y	68	Knee ext HHD no fixation/biodex	0.87 (0.43-0.95)
Neil, S	10	Knee flex	0.75 (0.38-0.82)
		Knee ext	0.35 (0.0-0.72)
Mentiplay, B	30	DF Lafayette-Kin Com	0.62 (0.15-0.83)
		DF Hoggan-Kin Com	0.61 (0.09-0.83)
		PF Lafayette-Kin Com	0.51 (0.12-0.78)
		PF Hoggan-Kin Com	0.47(0.0-0.78)
		Knee flex Lafayette-Kin Com	0.94 (0.87-0.97)
		Knee flex Hoggan-Kin Com	0.94(0.85-0.97)
		Knee ext Lafayette-Kin Com	0.82 (0.52-0.92)
		Knee ext Hoggan-Kin Com	0.90 (0.76-0.96)
		Hip abd Lafayette-Kin Com	0.88 (0.74-0.95)
		Hip abd Hoggan-Kin Com	0.89 (0.75-0.95)
		Hip add Lafayette-Kin Com	0.95 (0.87-0.98)
		Hip add Hoggan-Kin Com	0.94 (0.84-0.98)
		Hip flex Lafayette-Kin Com	0.94 (0.87-0.97)
Hip flex Hoggan-Kin Com	0.94 (0.85-0.97)		
Hip ext Lafayette-Kin Com	0.88 (0.72-0.95)		
Hip ext Hoggan-Kin Com	0.90 (0.76-0.95)		

Abbreviations: abd, abduction; add, adduction; DF, dorsiflexion; ext, extension; flex, flexion; HHD, hand-held dynamometry; ID, isokinetic dynamometry; ICC inter-rater correlation coefficient.

extension analysis [26,30-34]. All assessments were done while seated for knee flexion and extension.

Ankle plantar flexion and dorsiflexion. Two studies were included for ankle plantar flexion analysis [17,30]. The Rev 9000 [17], and Biodex [30] IDs were used. Assessments were performed in either a seated [17] or supine position [30]. Additionally, four studies were included for ankle dorsiflexion analysis [17,26,30,38]. These studies used the Kin Com [17,26] and Biodex [30,38] devices. Assessments were performed while either seated [17,38] or in a supine position [26,30].

4.3.3 Concurrent Validity between HHD and ID (Table 9)

Hip Joint

Two studies were included for inter-device ICC analysis in the hip (i.e. adduction, abduction, flexion, extension). In all cases, one study compared the Lafayette HHD to the Biodex ID [26] while the second study compared the Lafayette and Hoggan Health HHDs to the Kin Com device [29].

Table 5: Methodological quality by COSMIN checklist of the studies analyzing absolute reliability of HHD and ID.

Author	Was the percentage of missing values given?	Was there a description of how missing values were handled?	Was the sample size analysis adequate?	Important flaws Random allocation	Were at least two measurements available?	Were the administrations independent?	Was the time interval stated?	Was the time interval appropriate?	Were the test conditions similar for both measurements?	Was the SEM; SDC, or LOA calculated?	Rating
Arnold, C	Good	Fair	Poor	Exc	Exc	Fair	Exc	Fair	Exc	Exc	Poor
Kelln, B	Good	Fair	Poor	Exc	Exc	Fair	Exc	Fair	Exc	Exc	Poor
Kawaguchi, J	Good	Fair	Poor	Exc	Poor	Exc	Exc	Poor	Exc	Exc	Poor
Mentiplay, B	Good	Fair	Fair	Exc	Poor	Exc	Exc	Poor	Exc	Exc	Poor
Tung Wu, Lu	Poor	Fair	Poor	Exc	Exc	Exc	Exc	Exc	Exc	Exc	Poor
Claiborne, T	Good	Fair	Poor	Poor	Exc	Poor	Exc	Exc	Exc	Exc	Poor
Dauty, M	Good	Fair	Poor	Exc	Exc	Exc	Exc	Exc	Exc	Exc	Poor
Hartmann, A	Good	Fair	Poor	Exc	Exc	Exc	Exc	Fair	Exc	Exc	Poor
Pereira de Carvalho, A	Good	Fair	Poor	Poor	Exc	Poor	Exc	Exc	Exc	Exc	Poor
Ferri Morales, A	Good	Fair	Fair	NA	Good	NA	Exc	Good	Exc	Exc	Fair
Dervisevic, E	Good	Fair	Poor	Poor	Exc	Poor	Exc	Exc	Exc	Exc	Poor
Larsson, B	Good	Fair	Poor	NA	Good	NA	Exc	Good	Exc	Exc	Poor
Morrison, K	Good	Fair	Poor	Poor	Exc	Poor	Exc	Exc	Exc	Exc	Poor
Holmback, A	Good	Fair	Fair	NA	Exc	Exc	Exc	Exc	Exc	Exc	Fair
Kim, WK	Good	Fair	Fair	Exc	Exc	Exc	Exc	Exc	Exc	Exc	Fair

Abbreviations: Exc, excellent; NA, not applicable

Table 6: Methodological quality by COSMIN checklist of the studies analyzing concurrent validity between HHD and ID.

Author	Was the percentage of missing values given?	Was there a description of how missing values were handled?	Was the sample size analysis adequate?	Important flaws Random allocation	Can the employed criteria be considered on par with the gold standard?	Statistical methods	Rating
Arnold, C	Good	Fair	Poor	Excellent	Good	Excellent	Poor
Wang, Y	Good	Fair	Good	Excellent	Excellent	Excellent	Fair
Neil S,	Good	Fair	Poor	Fair	Excellent	Good	Poor
Mentiplay B,	Good	Fair	Fair	Excellent	Excellent	Good	Fair

Knee flexion

Two studies were included for inter-device ICC knee flexion analysis [26,36]. One study compared a PR1 HHD

to a Biodex ID [36]. The other included study compared the Lafayette and Hoggan Health HHDs with the Kin Com ID [26].

Table 7: Maximal voluntary isometric strength LOA for inter-subject variability between trials: hip, knee, and ankle muscles measured with HHD

Joint	Movements	Studies	Participants	SEM (%) lower and upper limit of studies	Statistical Method (%)	I ² (%)	P value	Effect size Kg (%)
Hip	Abductors	4	76	1.8, 14.1	LOA	0	0.839	12.78 (8.17, 17.38)
	Adductors	2	50	2.3, 6.8	LOA	0	0.54	12.04 (6.36, 17.72)
	Flexors	3	68	0.8, 21.1	LOA	67.3	0.02	22.29 (9.26, 35.32)
	Extensors	4	76	4.8, 18.1	LOA	0	0.77	19.56 (11.27, 27.85)
Knee	Flexors	3	75	3.8, 13.5	LOA	0	0.65	16.92 (8.61, 25.24)
	Extensors	5	123	4.0, 14.7	LOA	22	0.27	33.59(23.91, 43.26)
Ankle	Plantar Flexion	3	73	6.1, 8.5	LOA	0	0.9	48.87(35.19, 62.56)
	Dorsiflexion	3	68	2.2, 14.9	LOA	0	0.85	20.19(14.66, 25.72)

Table 8: Maximal voluntary isometric strength LOA for Inter-subject variability between trials: hip, knee, and ankle muscles measured with ID.

Joint	Movements	Studies	Participants	SEM (%) lower and upper limit of studies	Statistical Method (%)	I ² (%)	P value	Effect size Nm (%)
Hip	Abductors	2	53	6.2, 11.0	LOA	0	0.85	7.14 (-7.76, 22.03)
	Adductors	2	53	4.5, 18.8	LOA	0	0.75	0.99 (-11.1, 13.06)
	Flexors	2	53	6.5, 10.8	LOA	0	0.52	9.99 (-5.33, 25.31)
	Extensors	2	53	7.0, 7.4	LOA	0	0.94	1.09 (-13.05, 15.24)
Knee	Flexors	3	100	4.9, 15.9	LOA	0	0.9	2.09 (-6.4, 10.6)
	Extensors	5	150	3.6, 10.9	LOA	0	0.99	3.32 (-2.62, 9.26)
Ankle	Plantar Flexion	2	50	5.8, 13.8	LOA	0	0.89	1.75 (-26.4, 27.8)
	Dorsiflexion	3	80	5.9, 10.9	LOA	0	0.86	2.73 (-9.6, 15.1)

Table 9: Concurrent validity measured by ICC between HHD and isokinetic dynamometry

Joint	Movements	Studies	Participants	Statistical Method	I ² (%)	P value	Effect size
HIP	Abductors	2	48	ICC and 95% CI	0	0.42	0.87(0.80,0.94)
	Adductors	2	48	ICC and 95% CI	0	0.79	0.94(0.91,0.98)
	Flexors	2	48	ICC and 95% CI	32.8	0.23	0.93(0.88,0.98)
	Extensors	2	48	ICC and 95% CI	52.7	0.12	0.85(0.73,0.97)
Knee	Flexors	2	40	ICC and 95% CI	0	0.9	0.78(0.66,0.91)
	Extensors	5	126	ICC and 95% CI	80	0.0	0.64(0.42,0.85)
Ankle	Plantar flexion	2	40	ICC and 95% CI	12.5	0.31	0.63(0.45,0.81)
	Dorsiflexion	1	30	ICC and 95% CI	0	0.96	0.62(0.37,0.87)

Knee extension

Four studies were included for inter-device ICC knee extension analysis [11,26,35,36]. In one study, the Lafayette HHD was compared to the Biodex ID [11]. Another of the studies compared the Lafayette and Hoggan Health HHDs with the Kin Com ID [26]. In turn, the third study compared and Integrated Load Cell HHD with the Biodex ID [35] while the final report compared the PR1 HHD to the Biodex ID [36].

A single study was included for inter-device ICC analyses in the ankle (i.e. plantar and dorsiflexion) comparing Lafayette and Hoggan Health HHDs to the Kin Com device [29].

5 Discussion

Researchers and clinicians should recommend the use of strength assessment tools that have < 15% LOA between trials [39] to be able to detect small but clinically relevant changes in strength assessment. Based on the results of this systematic review, this goal is difficult to attain not only for HHD systems, but also for IDs. Although isokinetic dynamometry is the current gold standard for strength assessments [15], it was expected that the upper LOA limit of the assessed studies would fall much closer to 15%. Instead, this only occurred for a few movements and joints (i.e. hip adduction and extension; knee flexion and extension; and ankle dorsiflexion). Nevertheless, hip, knee, and ankle assessments did show a lower LOA when using IDs as compared to HHDs. While IDs are much more expensive and not portable, these devices remain more reliable than HHDs in strength assessments. Therefore, we recommend that IDs be used for these joints and movements.

For all HHD assessments of different joints, the upper LOA limit was always greater than 15%. This means that for any future value to be considered outside the range of random instrument error, differences would have to be 15% higher or lower than initial values. This is problematic in that muscle force improvements or deteriorations < 15% are still considered clinically relevant but would not be detected by HHDs [40]. Of all the joints analyzed, higher LOAs and, consequently, lower reliability were found for assessments of knee extension and ankle plantar flexion.

It is worth noting that several factors can affect the reliability of strength-related assessment tools [5-7]. These factors will be described below.

Position

Body position when performing a strength assessment is relevant for results. For example, Edouard et al [41] reported that the reliability of ID shoulder rotator strength assessments is dependent on shoulder position (i.e. frontal or scapular plane with 45° or 90° of abduction). When evaluating commonly referenced texts on the topic of manual muscle testing [42,43], it becomes clear that there is a fundamental lack of consensus for patient and practitioner positions. Nevertheless, when standardized techniques are used, the inter- and intra-rater reliability of manual testing prominently improves in healthy populations [44,45].

The results of the present review further highlight the lack of consistency for position when assessing muscle strength. The hip abduction studies, for example, varied between standing [11,25] and a supine position [9,26]. Likewise, knee flexion was assessed while either seated [26,27] or in a prone position [9]. When evaluating the reliability of HHDs, these different positions may reduce result accuracy. The highest heterogeneity ($I^2 = 67.3$) and upper LOA limit (35.2%) for hand-held dynamometry was found for hip flexion assessments, as compared to other hip movements. This could be influenced by the different assessment positions (e.g. standing [11], supine [9] and seated [26]), as compared to hip extension assessments, which only used a prone position. The upper LOA limits for knee flexion and extension were only 10.6% and 9.3%, respectively, showing accuracy in measuring peak torque, as compared with IDs. It is probable that this lower LOA was obtained due to all assessed studies using the same sitting position, as well as to IDs having good stabilization systems for assessments in this joint.

Evaluator and muscle group strengths

The reliability of the HHD test is known to increase when the rater is stronger than the subject [1]. For example, the knee extensors are a very strong muscle group and require a strong practitioner for accurate testing [13,46]. This is congruent with the higher LOAs observed in HHDs testing knee extension (43.3%) and ankle plantar flexion (62.6%), as compared to the LOAs observed in IDs for knee extension (9.3%) and ankle plantar flexion (23.8%), where evaluator muscle force is not a factor (Tables 7 and 8). These inter-device differences were the highest among all the joints and movements included for assessment, demonstrating that muscle group strength is one of the most

relevant factors affecting reliability in strength-related studies.

Fixation system

The strap system applied for body fixation and patient comfort during testing [47] may contribute noise to the measurements that can alter subject performance. Burnham [48] published that if the patient is not well stabilized during HHD testing, other muscles will be involved with the process and affect reliability. HHDs do not have their own stabilization straps, and the stabilization procedure is usually unclear or not described at all in studies. Alfuth [49] reported fair ICC values (0.58, 0.82) for ankle inversion and moderate values (0.77, 0.87) for ankle eversion using a HHD. Although these are not strong muscle groups, difficulties in stabilizing the ankle joint may have contributed to this low reliability. The present systemic review found that the highest upper LOA limits using ID were for hip abduction (22.03%) and flexion (25.31%), as well as for ankle plantar flexion (23.89%) (Table 8). Difficulties in stabilizing the lumbopelvic region in hip assessments and the ankle may also be important factors affecting absolute reliability when using an ID. It is likely that if IDs would have better stabilization systems for these strength assessments, lower LOA values would be obtained.

Devices

There is still controversy on the reliability between isokinetic devices. While Thompson [50] suggested no difference between the Biodex and Cybex IDs for the knee flexors, Gross [51] demonstrated that knee flexion tests performed on the Cybex ID reached higher peak torque values.

Biological factors

While it could be possible to control all of the aforementioned factors, variations between measurements would still exist for each subject due to biological factors. This is the result of changes in mental or physical states between trials, which is equally applicable for the tester or the person assessed [40].

Limits of agreement for decision making

The LOA is a stringent decision limit for establishing improvement/deterioration in peak muscle force or torque following rehabilitation post-injury or as part of a strengthening program in a healthy individual. High heterogeneity between subjects exists for many measurements in sports medicine, as in the case of peak muscle force. Therefore, SEMs and, consequently, LOAs are high [52]. Experts in sports medicine rehabilitation, consider 10% to be a clinically relevant improvement or deterioration in muscle force [39]. In practice, one criterion for a return to sports is peak muscle strength deficits under 10% of the contralateral extremity. This small, but clinically relevant, difference was only detected during knee flexion (-6.43, 10.61%) and knee extension (-2.62, 9.26) using IDs. All remaining assessments using an ID, as well as full strength assessments using HHDs, showed LOA values higher than the desired 10%. As previously mentioned, this means that these devices are not able to accurately detect these clinically meaningful changes.

Concurrent validity (ICC) between HHDs and IDs

A very high correlation was found for the hip adductors and flexors. A high correlation was found for the hip abductors and extensors, and a moderate correlation existed for the knee flexors and extensors, as well as ankle plantar and dorsiflexion. The highest CI was recorded for the knee extensors and ankle plantar and dorsiflexion. Lower concurrent validity and wider CI appears in joints and movements that show higher LOAs assessed by HHDs.

Methodological elements

Considering the Cosmin checklist [22], the included studies were classified from fair and poor methodologically. Final Cosmin classifications are determined by the lowest score in any of the analyzed items. As a result, only three studies [16,24,32] ranked as fair in relation to absolute reliability, while the remaining were scored as poor. The major methodological flaw of these studies was related to sample size. To obtain a score of good, at least 50 subjects should be recruited. This sample size was not achieved by any study. Additionally, no study indicated if dropouts existed or how data loss was managed. Another crucially important item is the independence of measurements. Terwee *et al* [37] recommends at least one week

between trials to ensure independence, a time met by only 8 studies.

Four studies [11,26,35,36] reviewed the concurrent validity between HHDs and IDs, as expressed through an inter-device ICC correlation. Two of them [11,36] were scored as poor because of a low sample size (< 20). One study [26] was found methodologically fair, [26] while the last was good. [35] The gold standard was always recognized or assumed.

Publication bias

A potential bias is the omission of non- English publications. There was no attempt to identify unpublished studies and doctoral thesis in this area.

Strengths and limitations

To our knowledge, this is the first systematic review and meta-analysis focusing on the absolute reliability and concurrent validity of HHDs and IDs. A comprehensive search was performed for all relevant published research over a wide range of years (1965-2016). There was restricted information regarding concurrent validity due to the scarcity of studies providing adequate data, and thus, the generalizability of the results is limited. Finally, the quality of the included studies rated from mostly poor to fair, which impeded subgroup analyses based on quality.

6 Conclusions

Considering COSMIN classifications, the assessed studies ranked methodologically between fair and poor. Considering all HHD assessments, the highest LOAs and, therefore, lower reliability scores were found for knee extension and ankle plantar flexion. We therefore suggest that another instrument be used to assess the peak torque of these movements. Considering all ID assessments, only hip adduction and extension; knee flexion and extension; and ankle dorsiflexion showed LOAs close to 15%. Hip, knee, and ankle assessments showed lower LOAs when using an ID compared to an HHD. A very high correlation was found for the hip adductors and flexors. In turn, a high correlation was found for the hip abductors and extensors, while a moderate correlation existed for the knee flexors and extensors, as well as ankle plantar and dorsiflexion.

Conflict of interest statement: Authors state no conflict of interest

References

- [1] Deones V.L., Wiley S.C., Worrell T., Assessment of quadriceps muscle performance by a hand-held dynamometer and an isokinetic dynamometer, *J. Orthop. Sports Phys. Ther.*, 1994, 20, 296-301
- [2] Trudelle-Jackson E., Jackson A.W., Frankowski C.M., Long K.M., Meske N.B., Interdevice reliability and validity assessment of the Nicholas Hand-Held Dynamometer, *J. Orthop. Sports Phys. Ther.*, 1994, 20, 302-306
- [3] Li R.C., Jasiewicz J.M., Middleton J., Condie P., Barriskill A., The development, validity, and reliability of a manual muscle testing device with integrated limb position sensors, *Arch. of Phys. Med. Rehabil.*, 2006, 87, 411-417
- [4] Mayer F., Horstmann T., Kranenberg U., Röcker K., Dickhuth H., Reproducibility of isokinetic peak torque and angle at peak torque in the shoulder joint, *Int. J. Sports Med.*, 1994, 15, S26-31
- [5] Kimura I.F., Gulick D.T., Alexander D.M., Takao S.H., Reliability of peak torque values for concentric and eccentric shoulder internal and external rotation on the Biodex, Kinetic Communicator, and Lido dynamometers, *Isokinet. Exerc. Sci.*, 1996, 6, 95-99
- [6] Soderberg G.J., Blaschak M., Shoulder internal and external rotation peak torque production through a velocity spectrum in differing positions, *J. Orthop. Sports Phys. Ther.*, 1987, 8, 518-524
- [7] Walmsley R.P., Szybbo C., A comparative study of the torque generated by the shoulder internal and external rotator muscles in different positions and at varying speeds, *J. Orthop. Sports Phys. Ther.*, 1987, 9, 217-222
- [8] Stark T., Walker B., Phillips J.K., Fejer R., Beck R., Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review, *Pm & R*, 2011, 3, 472-479
- [9] Kelln B.M., McKeon P.O., Gontkof L.M., Hertel J., Hand-held dynamometry: reliability of lower extremity muscle testing in healthy, physically active, young adults, *J. Sport Rehabil.*, 2008, 17, 160-170
- [10] Lu T., Hsu H., Chang L., Chen H., Enhancing the examiner's resisting force improves the reliability of manual muscle strength measurements: comparison of a new device with hand-held dynamometry, *J. Rehabil. Med.*, 2007, 39, 679-684
- [11] Arnold C.M., Warkentin K.D., Chilibeck P.D., Magnus C.R., The reliability and validity of handheld dynamometry for the measurement of lower-extremity muscle strength in older adults, *J. Strength Cond. Res.*, 2010, 24, 815-824
- [12] Campos Jara C.A., Bautista González I.J., Chiroso Ríos L.J., Martín Tamayo I., Lopez Fuenzalida A.E., Chiroso Ríos I.J., Validación y fiabilidad del dispositivo Haefni Health System 1.0 en la medición de la velocidad en el rango isocinético, *Cuadernos de Psicología del Deporte*, 2014, 14, 91-98
- [13] Martin H., Yule V., Syddall H., Dennison E., Cooper C., 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, 154-159
- [14] Meyer C., Corten K., Wesseling M., Peers K., Simon J.P., Test-retest reliability of innovated strength tests for hip muscles. *PLoS ONE [Electronic Resource]*, 2013, 8, e81149
- [15] Drouin J.M., Valovich-mcLeod T.C., Shultz S.J., Gansneder B.M., Perrin D.H., Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements, *Eur. J. Appl. Physiol.*, 2004, 91, 22-29
- [16] Holmbäck A.M., Lexell J., Reproducibility of isokinetic ankle dorsiflexor strength and fatigue measurements in healthy older subjects, *Isokinet. Exerc. Sci.*, 2007, 15: 263-270
- [17] Morrison K.E., Kaminski T.W., The reproducibility of an isokinetic testing technique at the ankle joint, *Isokinet. Exerc. Sci.*, 2007, 15, 245-251
- [18] Liberati A., Altman D.G., Tetzlaff J., Mulrow C., Gotzsche P.C., The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration, *Ann. Intern. Med.*, 2009, 151, W-65-W-94
- [19] Stratford P.W., Getting more from the literature: estimating the standard error of measurement from reliability studies, *Physiother. Can.*, 2004, 56, 27-30
- [20] Stratford P.W., Goldsmith C.H., Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Phys. Ther.*, 199, 77, 745-750
- [21] Golriz S., Hebert J.J., Foreman K.B., Walker B.F., The validity of a portable clinical force plate in assessment of static postural control: concurrent validity study, *Chiropr. Manual Ther.*, 2012, 20, 1-8
- [22] Terwee C.B., Mokkink L.B., Knol D.L., Ostelo R.W., Bouter L.M., Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist, *Qual. Life Res.*, 2012, 21, 651-657
- [23] Munro B.H., *Statistical methods for health care research*, Lippincott Williams & Wilkins, 2005
- [24] Kim W.K., Kim D.K., Seo K.M., Kang S.H., 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, 84-93
- [25] Kawaguchi J.K., Babcock G., Validity and reliability of handheld dynamometric strength assessment of hip extensor and abductor muscles, *Athletic Training and Sports Health Care*, 2010, 2, 11-17
- [26] Mentiplay B.F., Perraton L.G., Bower K.J., Adair B., Pua Y.H., Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: A reliability and validity study, *PloS one*, 2015, 10, e0140822
- [27] Tung-Wu L., Hui-Lien C., Ling-Ying C., Horng-Chaung H., Enhancing the examiner's resisting force improves the validity of manual muscle strength measurements: application to knee extensors and flexors, *J. Strength Cond. Res.*, 2012, 26, 2364-2371
- [28] Claiborne T.L., Timmons M.K., Pincivero D.M., Test-retest reliability of cardinal plane isokinetic hip torque and EMG, *J. Electromyogr. Kinesiol.*, 2009, 19, e345-352
- [29] Dauty M., Rochcongar P., Reproducibility of concentric and eccentric isokinetic strength of the knee flexors in elite volleyball players, *Isokinet. Exerc. Sci.*, 2001, 9: 129-132
- [30] Hartmann A., Knols R., Murer K., de Bruin E.D., Reproducibility of an isokinetic strength-testing protocol of the knee and ankle in older adults, *Gerontology*, 2009, 55, 259-268
- [31] de Carvalho F., Andrade A.C., Caserotti P., de Carvalho C.M., de Azevedo Abade E.A., da Eira Sampaio A.J., Reliability of concentric, eccentric and isometric Knee extension and flexion when using the REV9000 isokinetic dynamometer, *J. Hum. Kinet.*, 2013, 37, 47-53
- [32] Ferri-Morales A., Alegre L., Basco A., Aguado X., Test-retest relative and absolute reliability of knee extensor strength measures and minimal detectable change, *Isokinet. Exerc. Sci.*, 2014, 22, 17-26
- [33] Dervisevic E., Hadzic V., Karpljuk D., Radjio I., The influence of different ranges of motion testing on the isokinetic strength of the quadriceps and hamstrings, *Isokinet. Exerc. Sci.*, 2006, 14, 269-278
- [34] Larsson B., Karlsson S., Eriksson A., Gerdle B., Test-retest reliability of EMG and peak torque during repetitive maximum concentric knee extensions, *J. Electromyogr. Kines.*, 2003, 13, 281-287
- [35] Wang Y-C., Bohannon R.W., Magasi S.R., Hrynkiwicz B., Morales A., Testing of knee extension muscle strength: A comparison of two portable alternatives for the NIH toolbox study, *Isokinet. Exerc. Sci.*, 2011, 19, 163-168
- [36] Neil S.E., Myring A., Peeters M.J., Pirie I., Jacobs R., Reliability and validity of the Performance Recorder 1 for measuring isometric knee flexor and extensor strength, *Physiother. Theory & Pract.*, 2003, 29, 639-647
- [37] Terwee C.B., Bot S.D., de Boer M.R., van der Windt D.A., Knol D.L., Quality criteria were proposed for measurement properties of health status questionnaires, *J. Clin. Epidemiol.*, 2007, 60, 34-42
- [38] Holmbäck A.M., Porter M.M., Downham D., Lexell J., Reliability of isokinetic ankle dorsiflexor strength measurements in healthy young men and women, *Scan J. Rehabil. Med.*, 1999, 31, 229-239
- [39] Prentice W.E., Kaminski T.W., *Rehabilitation techniques for sports medicine and athletic training*, McGraw-hill, New York, 2004
- [40] Hopkins W.G., *Measures of reliability in sports medicine and science*, *Sports medicine*, 2000, 30, 1-15
- [41] Edouard P., Codine P., Samozino P., Bernard P.L., Hérisson C., Reliability of shoulder rotators isokinetic strength imbalance measured using the Biodex dynamometer, *J. Sci. Med. Sport*, 2013, 16, 162-165
- [42] Hislop H., Avers D., Brown M., Daniels and Worthingham's muscle testing: Techniques of manual examination and performance testing, Elsevier Health Sciences, 2013
- [43] Reese N.B., *Muscle and sensory testing*, Elsevier Health Sciences, 2013
- [44] Frese E., Brown M., Norton B.J., Clinical reliability of manual muscle testing, *Phys. Ther.*, 1987, 67, 1072-1076
- [45] Brandsma J.W., Schreuders T.A., Birke J.A., Piefer A., Oostendorp R., Manual muscle strength testing: intraobserver and interobserver reliabilities for the intrinsic muscles of the hand, *J. Hand Ther.*, 1995, 8, 185-190
- [46] Bohannon R.W., Hand-held compared with isokinetic dynamometry for measurement of static knee extension torque (parallel reliability of dynamometers), *Clin. Phys. Physiol. Meas.*, 1990, 11, 217-222

- [47] de Araujo Ribeiro Alvares J.B., Rodrigues R., de Azevedo Franke R., da Silva B.G., Pinto R.S., Inter-machine reliability of the Biodex and Cybex isokinetic dynamometers for knee flexor/ extensor isometric, concentric and eccentric tests, *Phys. Ther. Sport*, 2015, 16, 59-65
- [48] Burnham R.S., Bell G., Olenik L., Reid D.C., Shoulder abduction strength measurement in football players: reliability and validity of two field tests, *Clin. J. Sport Med.*, 1995, 5, 90-94
- [49] Alfuth M., Hahm M.M., Reliability, comparability, and validity of foot inversion and eversion strength measurements using a hand-held dynamometer, *Int. J. Sports Phys. Ther.*, 2016, 11, 72-79
- [50] Thompson M.C., Shingleton L.G., Kegerreis S.T., Comparison of values generated during testing of the knee using the Cybex II Plus® and Biodex Model B-2000® isokinetic dynamometers, *J. Orthop. Sports Phys. Ther.*, 1989, 11, 108-115
- [51] Gross M.T., Huffman G.M., Phillips C.N., Wray J.A., Intramachine and intermachine reliability of the Biodex and Cybex® II for knee flexion and extension peak torque and angular work, *J. Orthop. Sports Phys. Ther.*, 1991, 13: 329-335
- [52] Nevill A.M., Atkinson G., Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science, *Br. J. Sports Med.*, 1997, 3