UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Reliability and criterion validity of handheld dynamometry for measuring trunk muscle strength in people with and without chronic non-specific low back pain

Althobaiti, Shouq; Falla, Deborah

DOI:
10.1016/j.msksp.2023.102799

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Althobaiti, S & Falla, D 2023, 'Reliability and criterion validity of handheld dynamometry for measuring trunk muscle strength in people with and without chronic non-specific low back pain', *Musculoskeletal Science and Practice*, vol. 66, 102799. https://doi.org/10.1016/j.msksp.2023.102799

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 18. Sep. 2024

Musculoskeletal Science and Practice 66 (2023) 102799

Contents lists available at ScienceDirect

Musculoskeletal Science and Practice

journal homepage: www.elsevier.com/locate/msksp

Original article

Reliability and criterion validity of handheld dynamometry for measuring trunk muscle strength in people with and without chronic non-specific low back pain

Shouq Althobaiti [a,b,*], Deborah Falla [a]

[a] Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham, UK

[b] Physical Therapy Department, College of Applied Medical Science, Taif University, Taif, Saudi Arabia

ARTICLE INFO

Keywords:
Handheld dynamometer
Low back pain
Trunk muscle strength
Psychometric properties

ABSTRACT

Background: Evaluating trunk strength is an important aspect of the physical examination of people with low back pain (LBP). Thus, reliable, valid, and easily applied measurement tools are needed to quantify trunk muscle strength and monitor changes in response to interventions.

Objectives: To determine within-day and between-day test re-test reliability and criterion validity of a handheld dynamometer (HHD) to evaluate maximum isometric trunk strength in people with chronic LBP and asymptomatic individuals.

Design: Reliability and criterion validity study.

Methods: Twenty adult participants with chronic, non-specific LBP and 35 asymptomatic individuals participated. Isometric trunk flexion, extension, and rotation strength were evaluated with the HHD (Active force 2) and the within-day and between-day reliability were determined with intraclass correlation coefficients ($ICC_{2,1}$) and the standard error of the measurements (SEM), minimal detectable change (MDC), and the limits of agreement (LOA) using Bland-Altman plots. Criterion validity was evaluated using Pearson correlation coefficients to compare HHD measurements to isokinetic dynamometry for both isometric trunk flexion and extension strength.

Results: Good to excellent within-day and between-day reliability was observed for people with LBP and asymptomatic individuals with ($ICC_{2,1}$) of 0.73–0.93 and 0.62–0.92 respectively. A moderate to strong correlation was found between measurements with the HHD and the isokinetic dynamometer with a correlation of $r = 0.68–0.78$ and $r = 0.56–0.59$ for people with LBP and asymptomatic participants respectively.

Conclusion: A HHD is a reliable, valid, and clinically applicable tool for the measurement of trunk strength in adults with and without chronic LBP.

1. Introduction

Low back pain (LBP) affects more than 80% of people at some point during their lifetime (Taylor et al., 2014), with the peak incidence occurring between 30 and 65 years (Hoy et al., 2010). LBP may become chronic and consequently people often develop long-term physical and psychological disability (Pengel et al., 2003; Menezes Costa et al., 2012), which in turn leads to a substantial socioeconomic burden (Hong et al., 2013). Although 10–15% of LBP episodes can be attributed to an underlying pathoanatomical cause, most chronic LBP is classified as non-specific LBP (Airaksinen et al., 2006).

Trunk muscle strength is required for everyday activities to provide support, movement control and minimize the load on the spine (Granacher et al., 2013; Ebenbichler et al., 2001). Reduced trunk muscle strength (Cho et al., 2014; Steele et al., 2014) and altered activity of the trunk muscles (Sanderson et al., 2019; Arvanitidis et al., 2021) have commonly been identified in people with chronic LBP. Therefore, the measurement of trunk muscle strength is an important aspect of a functional evaluation of patients with chronic LBP to monitor improvements in strength in response to training interventions (Steele et al., 2019). To confidently interpret the observed gain or loss in strength as actual change and not secondary to measurement error,

* Corresponding author. Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences. University of Birmingham. UK.

E-mail address: d.falla@bham.ac.uk (S. Althobaiti).

https://doi.org/10.1016/j.msksp.2023.102799
Received 6 February 2023; Received in revised form 4 May 2023; Accepted 9 June 2023
Available online 10 June 2023
2468-7812/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

outcome measures must exhibit an acceptable level of psychometric properties (De Vet et al., 2011).

There are different measurement methods used for trunk strength evaluation with varying levels of psychometric properties; isokinetic dynamometry is the gold standard (Althobaiti et al., 2022; Guilhem et al., 2014). The use of isokinetic devices to measure trunk strength is considered safe (den Hartog et al., 2010), reliable (Verbrugghe et al., 2019), and sensitive to detect muscle weakness (Langrana et al., 1984). However, the practicality of isokinetic devices in clinical settings and home-based management is challenging due to the operational complexity and associated space requirements and costs (Mavroidis et al., 2005; Stark et al., 2011). A hand-held dynamometer (HHD) on the other hand has been considered as an alternative objective measurement tool since the convenient size, portability, and low cost of the HHD justifies its widespread use in clinical practice (Stark et al., 2011).

Evaluation of trunk strength using a HHD has been examined among asymptomatic and athletic populations with flexion and extension strength being the most tested trunk movements. Previous research demonstrated good to excellent test re-test reliability of the HHD with the intraclass correlation coefficients (ICC) > 0.7 when the HHD was passively fixated (closed chain) or held by an examiner (open chain) to evaluate trunk strength from sitting, standing, and lying positions. (Harding et al., 2017; Jubany et al., 2015; Park et al., 2017; De Blaiser et al., 2018; Alshammari et al., 2022). However, these findings cannot be translated to people with chronic LBP, and it is relevant to note that the psychometric properties of HHD when evaluating trunk strength of people with LBP has not been adequately examined to date. The only identified study according to a recent systematic review (Althobaiti et al., 2022) was by Ozcan Kahraman et al. (2016) who investigated the intra-rater reliability of the HHD in people with chronic LBP aiming to develop a core stability assessment battery. People with chronic LBP with moderate disability were examined and the results of the study showed moderate to good reliability of the HHD with an ICC = 0.65 and ICC = 0.74 for extension and flexion strength respectively (Ozcan Kahraman et al., 2016). The validity of HHD for the evaluation of trunk strength has not been tested.

The present study aimed to determine both within-day and between-day test re-test reliability and the criterion validity of a HHD for the measurement of trunk strength in people with and without chronic LBP.

2. Materials and methods

2.1. Study design and setting

This study evaluated within and between-day test-retest reliability and criterion validity of a HHD for trunk strength assessment in people with chronic LBP and asymptomatic controls. According to COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) definition, reliability refers to the extent to which the measurement instrument is devoid of any errors that could affect its accuracy, and criterion validity is defined as the degree to which the scores obtained from the instrument of interest are an adequate reflection of the gold standard (Mokkink et al., 2010).

Data were collected at a laboratory within the Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), University of Birmingham, UK, between September 2021 and January 2022. The reporting of this study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Von Elm et al., 2007). Ethical clearance for the study was obtained from the Ethics Committee of the University of Birmingham (ERN_21–0846), and all participants signed an informed consent form prior to data collection. All methods were performed in accordance with the Declaration of Helsinki.

2.2. Participants

Based on the power calculation of 0.05 significance level, with true reliability above 0.7 the required sample of ≥19 participants were needed (Walter et al., 1998). Regarding the validity study, a sample of 13 or more is needed to detect a correlation of 0.7, with α level of 0.5 and power of 80% (Faul et al., 2009). Considering the possible dropouts, a total of twenty adults with a history of chronic, non-specific LBP and 35 controls were recruited. The inclusion criteria for the LBP group were pain for at least 3 months (chronic) and score of at least mild pain intensity (equal to or greater than 2 out of 10) on a numerical rating scale (NRS). Where people with no history of LBP were included for control group. Exclusion criteria for both groups were history of systematic disease, and trunk or pelvic pain attributed to a specific cause (e.g., pain related to trauma, surgery, fractures, spinal stenosis, or tumour). For those with LBP, individuals under active management for their pain were also excluded.

2.3. Questionnaires

Following written informed consent, pain intensity of those with LBP was measured using the NRS, disability was measured using the Oswestry Disability Index (ODI), and the Tampa Scale for Kinesiophobia (TSK) was used to measure fear of movement. Health status and physical activity were measured using the 36-Item Short Form Survey (SF-36) and the International Physical Activity Questionnaire (IPAQ) respectively. Asymptomatic controls were asked to only complete the TSK and IPAQ. The questionnaires used are widely used in research and in clinical practice, and they have established psychometric properties (Beaton et al., 1997; Swinkels-Meewisse et al., 2003; Childs et al., 2005; Craig et al., 2003; Jenkinson et al., 1994).

2.4. Dynamometry

An Active Force 2 HHD (Active body Inc., United States) which was externally fixed, was used to measure maximum voluntary isometric contractions (MVIC) of trunk flexion, extension, and rotation (inner and outer range). Each task was measured from two different positions, except for rotation in those with chronic LBP which was measured in sitting only (in total 6 tasks for people with LBP and 8 tasks for asymptomatic participants were measured). The selected testing positions were based on previous literature (Jubany et al., 2015; De Blaiser et al., 2018; Johnson et al., 2012). The measurements were conducted in a random order to minimize systematic error using a random number generator. The participant was comfortably positioned into the testing positions as illustrated in Fig. 1, and the HHD was secured with a strap just under the suprasternal notch for flexion, and at the level of T4 for extension strength measurements. Rotation strength was measured from sitting with the hips and knees flexed to 90° and feet off the floor, with a belt secured around the pelvis and proximal to the knees to limit the engagement of lower extremities. While holding the bar with an upright position, participants were directed to rotate their trunk and push against the HHD, on the most painful side for the LBP group and the dominant side for asymptomatic controls. Participants were instructed to push forward to measure inner rotation strength and to push backward for outer rotation strength. The same protocol was performed in standing with feet positioned shoulders width apart.

A Biodex 3 isokinetic dynamometer (Medical System Inc., United States) was used to evaluate the criterion validity of the HHD measures. The participants were positioned in a trunk flexion-extension attachment in a semi-standing position as per the manufacturer's recommendation (Inc., 2020). From the same position, the HHD was positioned over the sternum, inferior to the suprasternal notch, for flexion measurements and secured to the scapula roll for extension strength measurements.

2.5. Experimental procedure

To evaluate within-day and between-day reliability, participants

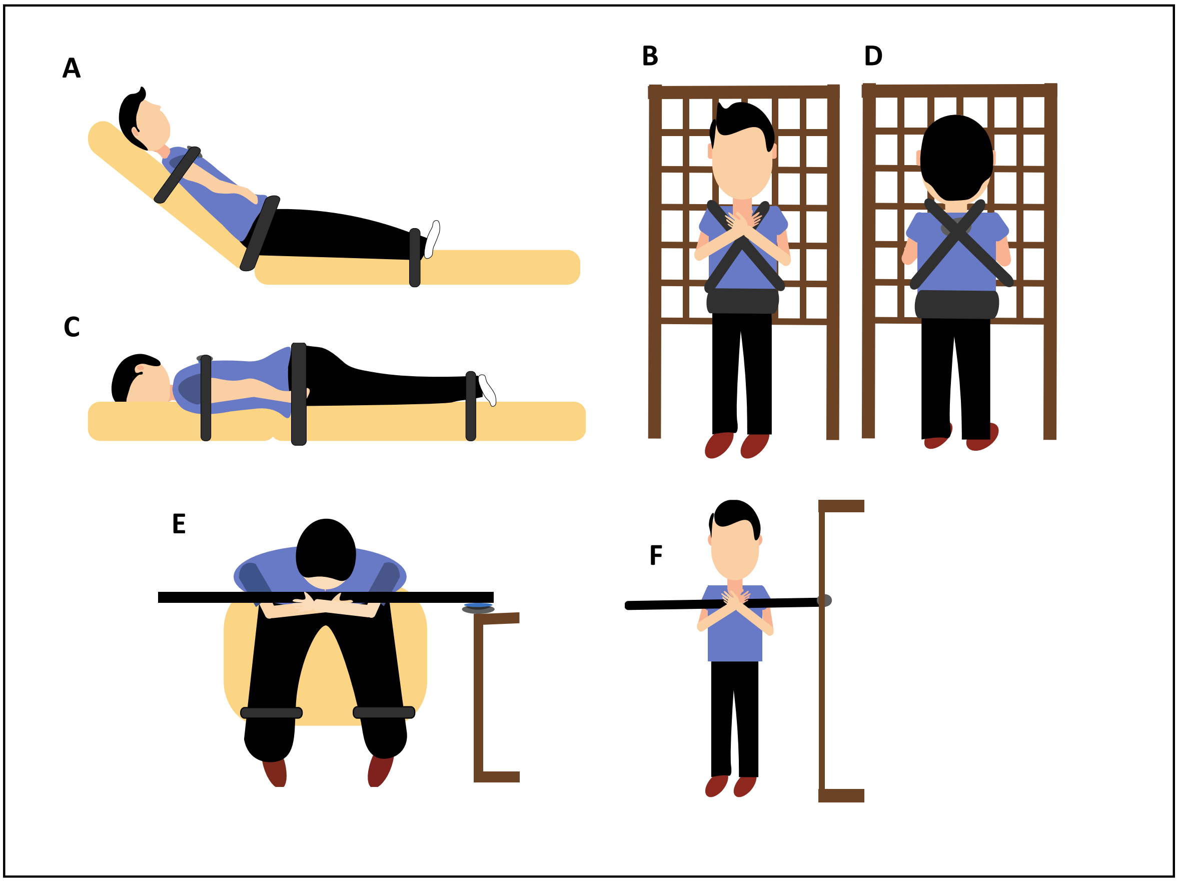

Fig. 1. Experimental setup for the isometric trunk strength testing using handheld dynamometry: Flexion **A)** 30° supported supine, **B)** standing. Extension **C)** prone, **D)** standing. Rotation **E)** sitting, **F)** standing.

attended four measurement sessions across two days. On day 1, sessions 1 and 2 utilized the HHD with a 30-minute rest period in between. On day 2, session 3 was conducted using the HHD to evaluate between-day reliability, and session 4 involved both the HHD and the isokinetic dynamometer.Participants were given a 5–10 days in between to allow for resolution of any muscle soreness but minimize any potential muscle adaptation between sessions (Demoulin et al., 2006). Participants were instructed to avoid any vigorous strength training directed to the trunk between measurement sessions. For the evaluation of criterion validity, only isometric trunk flexion and extension strength were measured using the Biodex 3 isokinetic dynamometer and HHD (order of testing randomised) with 15 min of rest between measures with the two devices.

All measurements were conducted by an experienced physiotherapist at the same time of the day for each participant to limit the influence of diurnal variations (Nuzzo et al., 2019). The participant started with 10 min of warming up on bicycle ergometer and prior to the actual measurements, 1–2 submaximal familiarization trials were performed (Pekünlü and Özsu, 2014), which was then followed by three trials of MVIC with 3–5 s hold. Two minutes rest was provided between trials and 3 min was provided between tasks to change position and allow for recovery (Haff, 2018; Gallagher et al., 2000).

Participants were requested to exert their maximum effort every trial and avoid holding their breath. Standardized testing conditions and instructions were provided as follows: *'I'm going to evaluate your trunk muscle strength, I need you to take a deep breath in while I'm counting down then breath out and push against the dynamometer gradually as hard as you can, keep pushing until I say relax'* for each trial, and additional verbal encouragement was given during the task. Participants were advised to report any pain or discomfort during the testing process.

2.6. Data analysis

The peak force measured in Newtons (N) for the HHD protocol and peak torque measured in Newton/meter for the isokinetic dynamometer were recorded and the highest value of the three attempts was taken for further analysis (Kato et al., 2019).

2.7. Statistical analysis

Data analysis was evaluated using IBM SPSS Statistics 28.0, with the normality of the data assessed using the Shapiro-Wilk test and the significance level was set at $p < 0.05$. Relative reliability for within day (sessions 1 & 2) and between days (sessions 2 & 3) was determined by ICC with a two-way random effects model, with absolute agreement, and single measurement ICC (Hoy et al., 2010; Taylor et al., 2014) (Trevethan, 2017). The ICC was interpreted as low with a value of <0.40, moderate 0.40–0.59, good 0.60–0.75, and excellent >0.75 (Cicchetti and Sparrow, 1981). As the relative reliability might be affected by the sample variability, the measurement error was also measured using the standard error of measurement (SEM) and limits of agreement (LOA), using Bland-Altman plots (Bland and Altman, 1986). Bland-Altman plots with 95% LOA were computed as follows: LOA = mean difference ± (1, 96 × $SD_{\text{mean difference}}$) for both comparisons (in this case between sessions 1 & 2 and sessions 2 & 3). One sample *t*-test of the difference was conducted to identify any trend of systematic error between sessions. The SEM was measured using the following formula: $SEM = \sqrt{\sigma 2_{\text{residual}}}$ (De Vet et al., 2011). The minimal detectable change was measured using the following formula: $MDC = SEM \times 1.96^{*}\sqrt{2}$.

Based on data normality, a two-tailed Pearson product-moment correlation coefficient (r) was used to evaluate the criterion validity between the isokinetic dynamometer and the HHD. The correlation was interpreted as low for scores <0.5, moderate 0.5–0.7, and strong >0.7.

3. Results

3.1. Participants' demographics and questionnaires

Twenty people with chronic LBP and 20 asymptomatic people participated in the reliability study. Another cohort of 15 asymptomatic participants was tested to establish the validity of the HHD together with the same 20 participants with LBP that participated in the reliability study. The baseline characteristics of the groups are presented in Table 1.

3.2. Reliability and measurement error

Overall, good to excellent reliability was observed for both within-day and between-day in asymptomatic participants and people with LBP (Tables 2 and 3). Regarding the LOA for between-day, a significant difference was found for trunk flexion strength from 30° supine in asymptomatic participants with a systematic bias of −12 N. In people with LBP, the Bland and Altman plot also shows a mean difference of −25 N and −29 N for extension and flexion strength from standing, which significantly differed from zero. All results are highlighted in **Appendix 1**.

3.3. Criterion validity

The relationship between the HHD and Biodex measures of trunk strength was examined using the Pearson correlation coefficient. A moderate correlation was observed in both groups (Table 4) which was statistically significant for those with LBP.

4. Discussion

The results revealed good to excellent within-day and between-day test re-test reliability for all tested trunk movements and in different testing positions. Additionally, a moderate to strong correlation was found between the HHD and the isokinetic dynamometer measures of trunk strength for flexion and extension measurements.

Previously published studies demonstrated sufficient levels of psychometric properties of the HHD for upper and lower limb strength measurement (Stark et al., 2011) and trunk strength measurement in chronic stroke patients (Karthikbabu and Chakrapani, 2017). The current findings are in line with previous research showing the reliability of isometric trunk muscle testing using HHD when tested in standing (Harding et al., 2017; Jubany et al., 2015), sitting (Park et al., 2017; Alshammari et al., 2022) and lying positions (De Blaiser et al., 2018; Tarca et al., 2020; Valentin and Maribo, 2014). On the contrary, a comparable study evaluating the measurement properties of the HHD

Table 2
Within day test re-test reliability.

	Task	ICC $_{2,1}$	95% CI for ICC	SEM (N)	MDC
Asymptomatic	Flexion, standing	0.82	0.61–0.92	20.03	55.52
	Flexion, 30° supine	0.78	0.53–0.90	14.70	40.76
	Extension, standing	0.88	0.73–0.95	24.99	69.27
	Extension, prone	0.78	0.53–0.90	19.56	54.23
	Inner rotation, standing	0.90	0.76–0.95	4.41	12.24
	Inner rotation, sitting	0.92	0.81–0.96	5.70	15.80
	Outer rotation, standing	0.68	0.37–0.86	7.96	22.07
	Outer rotation, sitting	0.85	0.65–0.93	7.16	19.86
LBP	Flexion, standing	0.93	0.84–0.97	20.15	55.87
	Flexion, 30° supine	0.80	0.55–0.91	15.64	43.36
	Extension, standing	0.88	0.72–0.95	26.29	72.88
	Extension, prone	0.78	0.53–0.90	16.7	46.31
	Inner rotation	0.91	0.80–0.96	7.42	20.57
	Outer rotation	0.92	0.82–0.96	6.88	19.08

ICC $_{2,1}$ = Intraclass correlation coefficient with two-way random effect model, CI= Confidence interval, SEM= Standard error of measurement, N= Newton, MDC = Minimal detectable change.

Table 1
Participants' baseline characteristics.

Characteristic	Reliability			Validity	
N	Asymptomatic Mean (SD) 20	LBP Mean (SD) 20	p- value	Asymptomatic Mean (SD) 15	p- value
Age (years)	25.6 (4.69)	26.6 (5.44)	0.42	28.47 (4.69)	0.32
Gender (Female %)	50%	60%		47%	
Weight (kg)	64.25 (12.59)	70.6 (13.12)	0.13	68.40 (13.33)	0.56
Hight (cm)	168.50 (11.71)	168.2 (10.76)	0.74	170.60 (8.71)	0.52
BMI (kg/cm^2)	22.78 (5.14)	24.9 (3.5)	0.22	23.44 (4.07)	0.35
NPRS (0–10)	-	4.66 (1.75)	-	-	-
ODI %	-	38.93 (6.79)	-	-	-
IPAQ					
Walking (min)	1856.8 (1841.38)	1237.5 (1037.64)	0.26	1288.1	0.84
Moderate (min)	739.16 (758.35)	1511.66 (2492.72)	0.31	716.81	0.46
Vigorous (min)	1156 (600.57)	2274.28 (1820.68)	0.16	1128	.15
TSK	31.46 (7.5)	37.93 (6.88)	**0.02****	28.13	**< 0.001****
SF-36					
Physical functioning	-	76.25 (13.62)	-	-	-
limitation due to physical health	-	43.75 (39.63)	-	-	-
limitation due to emotional problems	-	55 (39.40)	-	-	-
Energy/Fatigue	-	42.5 (13.72)	-	-	-
Emotional well-being	-	57.2 (19.08)	-	-	-
Social Functioning	-	66.25 (18.18)	-	-	-
Pain	-	57.38 (15.8)	-	-	-
General health	-	58.25 (12.7)	-	-	-

Kg = kilograms, BMI = body mass index, NPRS = numeric pain rating scale, ODI = Oswestry Disability Index, IPAQ = International Physical Activity Questionnaire, TSK = Tampa Scale for Kinesiophobia, SF-36 = 36-Item Short Form Survey, cm = centimetres, P value = paired-sample t-test.

Table 3
Between day test re-test reliability.

	Task	ICC $_{2,1}$	95% CI for ICC	SEM (N)	MDC
Asymptomatic	Flexion, standing	0.65	0.28–0.84	31.86	88.31
	Flexion, 30° supine	0.65	0.29–0.84	17.54	48.63
	Extension, standing	0.83	0.64–0.93	31.65	87.74
	Extension, prone	0.76	0.51–0.90	18.95	52.54
	Inner rotation, standing	0.62	0.27–0.83	7.27	20.15
	Inner rotation, sitting	0.79	0.56–0.91	8.75	24.26
	Outer rotation, standing	0.83	0.62–0.93	4.92	13.64
	Outer rotation, sitting	0.82	0.60–0.92	7.43	20.61
LBP	Flexion, standing	0.73	0.38–0.88	37.33	103.48
	Flexion, 30° supine	0.82	0.61–0.92	13.60	37.72
	Extension, standing	0.85	0.56–0.94	26.97	74.75
	Extension, prone	0.88	0.72–0.95	12.32	34.16
	Inner rotation	0.88	0.71–0.95	8.61	23.88
	Outer rotation	0.93	0.84–0.97	6.53	18.12

ICC $_{2,1}$ = Intraclass correlation coefficient with two-way random effect model, CI= Confidence interval, SEM= Standard error of measurement, N= Newton, MDC = Minimal detectable change.

Table 4
Criterion validity of the HHD.

	Task	r	95% CI		P value
			Lower bound	Upper bound	
Asymptomatic	Trunk flexion	0.56	.078	0.83	.027
	Trunk extension	0.59	.123	0.85	.019
LBP	Trunk flexion	0.78	0.52	0.90	<.001
	Trunk extension	0.68	0.34	0.86	<.001

has shown less reliable outcomes for trunk extension (ICC = 0.24) and flexion strength (ICC = 0.25) measurements (Moreland et al., 1997). This discrepancy in results is likely due to the testing position as the Sorensen test position was used in this earlier study and this position might be challenging due to pain in legs, abdomen and spine, especially for people with LBP (Biering-Sørensen, 1984).

Interestingly, the reliability results were superior for people with LBP compared to asymptomatic controls, and a possible explanation for that is that more tasks were required from the asymptomatic people compared to those with LBP, which may have resulted in fatigue thereby affecting the performance and subsequently the reliability of the measures.

As highlighted by Kumar (1997) studies evaluating trunk rotation performance are rare due to the lack of an objective, cost-effective and standardized assessment tool. The majority of the studies which have used HHD have only examined the measurement properties for flexion and extension strength measures (Althobaiti et al., 2022). The novel measurement protocol used in this study exhibited good to excellent reliability of rotation strength in both standing and sitting positions. These findings are comparable to reliability data obtained from other measurement devices (Andre et al., 2012; Kienbacher et al., 2016).

Trunk strength assessment may vary as the testing positions change (Kocjan and Sarabon, 2014). There are different testing protocols, and no standard position is recommended in the literature (Althobaiti et al., 2022). However, the adopted testing setup used in the current study was based on previous literature describing trunk MVIC measurement (Danneels et al., 2001; Konrad, 2005). The clinical feasibility and replicability of the testing positions were considered. In the current study, each movement was measured from two different positions, a more relaxed position (i.e., supported supine, prone and sitting) and a more functional position (i.e., standing), to mirror common everyday positions. A simple setup using only a treatment plinth, chair, wooden pole, and fixing belts was used in the current study to further facilitate the reproducibility of the current protocol and improve it is clinical applicability (afsluttede Den and Kandidatuddannelse, 2014).

The measurement error was evaluated using the SEM and the MDC, which can provide a direct evaluation of patient progress in clinical settings. It can be noted that the more stable positions such as 30° supine and prone for flexion and extension strength measures had lower measurement error compared to testing in standing. However, the systematic difference noted in the Bland and Altman plot can't be neglected as the flexion and extension strength measurements were overestimated in session 3 compared to session 2, suggesting a mild learning effect.

Regarding the criterion validity of the HHD, the findings are consistent with those of Yang and colleagues (Yang et al., 2020), where moderate correlations for extension strength were reported, with Pearson's coefficient ranging from 0.3 to 0.6. Although previously published studies have reported relatively higher validity between the HHD and isokinetic dynamometers (Harding et al., 2017; Jubany et al., 2015), this discrepancy could be due to the type of HHD used in the current study, which was slightly uncomfortable to push against, as reported by some participants. However, due to the measurement unit difference between both devices, it is still unclear whether this discomfort causes them to push less compared to when their strength is measured using the Biodex dynamometer.

4.1. Methodological considerations

Even though the influence of the examiner strength on the HHD measurements was eliminated with the use of external fixation, other sources of examiner-based variability such as instructions and participant positioning may have influenced the level of reliability (Kroemer and Marras, 1980). Additionally, due to logistical reasons, the investigator was not blind to the participants' performance which could have influenced the reliability and validity of the measures. However, measurements were taken, saved to the device, and then viewed by the rater only upon completion of all measurements.

The present work provides clinicians with reliable and valid trunk strength evaluation protocols to evaluate isometric trunk strength using a simple setup. However, data regarding the HHD predictive validity and sensitivity to the change in trunk strength is still lacking. Therefore, studies directed to investigate the predictive validity and responsiveness of the HHD for trunk strength measurement are also warranted. Given the characteristics of the current study sample, one limitation is the generalizability of the findings. In particular, older adults or people with higher levels of pain or disability were absent from our cohort. Although the sample size was sufficient to detect satisfactory test-retest reliability and validity levels, the COSMIN recommendations suggest that larger sample sizes are advisable. Consequently, further studies with larger sample sizes and more diverse participant characteristics are needed to enhance the generalizability of these findings. Although moderate agreement between the HHD and the isokinetic dynamometer was found, it is not recommended to use both methods interchangeably when evaluating trunk strength.

5. Conclusion

This study showed good to excellent within-day and between-day test re-test reliability of a HHD for measuring trunk strength, as well as moderate to strong correlation with measures obtained using an isokinetic dynamometer. Considering the current absolute reliability and measurement error, the use of 30° supported supine, prone, and sitting protocols is recommended for evaluating trunk muscle strength in people with chronic LBP.

Declaration of competing interest

None.

Acknowledgements

This work was funded by the Ministry of Education, Kingdom of Saudi Arabia via a scholarship awarded to SA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.msksp.2023.102799.

References

afsluttede Den, F.G.H.V., Kandidatuddannelse, S., 2014. Håndholdt Dynamometer.

Airaksinen, O., Brox, J.I., Cedraschi, C., Hildebrandt, J., Klaber-Moffett, J., Kovacs, F., et al., 2006. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur. Spine J. 15 (Suppl. 2), S192–S300. Suppl 2.

Alshammari, A., Algarni, N., Marwan, Y., Algarni, N., Weber, M., Ouellet, J., et al., 2022. Seated Lumbar Extensor Strength in Adults: a Study of Repeatability and Reliability.

Althobaiti, S., Rushton, A., Aldahas, A., Falla, D., Heneghan, N.R., 2022. Practicable performance-based outcome measures of trunk muscle strength and their measurement properties: a systematic review and narrative synthesis. PLoS One 17 (6), e0270101.

Andre, M.J., Fry, A.C., Heyrman, M.A., Hudy, A., Holt, B., Roberts, C., et al., 2012. A reliable method for assessing rotational power. J. Strength Condit Res. 26 (3), 720–724.

Arvanitidis, M., Bikinis, N., Petrakis, S., Gkioka, A., Tsimpolis, D., Falla, D., et al., 2021. Spatial distribution of lumbar erector spinae muscle activity in individuals with and without chronic low back pain during a dynamic isokinetic fatiguing task. Clin. BioMech. 81, 105214.

Beaton, D.E., Hogg-Johnson, S., Bombardier, C., 1997. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J. Clin. Epidemiol. 50 (1), 79–93.

Biering-Sørensen, F., 1984. Physical measurements as risk indicators for low-back trouble over a one-year period. Spine 9 (2), 106–119.

Bland, J.M., Altman, D., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327 (8476), 307–310.

Childs, J.D., Piva, S.R., Fritz, J.M., 2005. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine 30 (11), 1331–1334.

Cho, K.H., Beom, J.W., Lee, T.S., Lim, J.H., Lee, T.H., Yuk, J.H., 2014. Trunk muscles strength as a risk factor for nonspecific low back pain: a pilot study. Arm 38 (2), 234–240.

Cicchetti, D.V., Sparrow, S.A., 1981. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am. J. Ment. Defic. 86 (2), 127–137.

Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., et al., 2003. International physical activity questionnaire: 12-country reliability and validity. Med. Sci. Sports Exerc. 35 (8), 1381–1395.

Danneels, L.A., Vanderstraeten, G.G., Cambier, D.C., Witvrouw, E.E., Stevens, V.K., De Cuyper, H.J., 2001. A functional subdivision of hip, abdominal, and back muscles during asymmetric lifting. Spine 26 (6), E114–E121.

De Blaiser, C., De Ridder, R., Willems, T., Danneels, L., Roosen, P., 2018. Reliability and validity of trunk flexor and trunk extensor strength measurements using handheld dynamometry in a healthy athletic population. Phys. Ther. Sport 34, 180–186.

De Vet, H.C., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. Measurement in Medicine: a Practical Guide. Cambridge university press.

Demoulin, C., Grosdent, S., Debois, I., Mahieu, G., Maquet, D., Jidovsteff, B., et al., 2006. Inter-session, inter-tester and inter-site reproducibility of isometric trunk muscle strength measurements. Isokinet. Exerc. Sci. 14 (4), 317–325.

den Hartog, D., Eker, H., Tuinebreijer, W., Kleinrensink, G.J., Stam, H., Lange, J., 2010. Isokinetic strength of the trunk flexor muscles after surgical repair for incisional hernia. Hernia 14 (3), 243–247.

Ebenbichler, G.R., Oddsson, L.I., Kollmitzer, J., Erim, Z., 2001. Sensory-motor control of the lower back: implications for rehabilitation. Med. Sci. Sports Exerc. 33 (11), 1889–1898.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41 (4), 1149–1160.

Gallagher, S., Moore, J.S., Stobbe, T.J., McGlothlin, J.D., Bhattacharya, A., 2000. Physical Strength Assessment in Ergonomics. Handbook of Industrial Automation. CRC Press, pp. 797–827.

Granacher, U., Gollhofer, A., Hortobágyi, T., Kressig, R.W., Muehlbauer, T., 2013. The importance of trunk muscle strength for balance, functional performance, and fall prevention in seniors: a systematic review. Sports Med. 43 (7), 627–641.

Guilhem, G., Giroux, C., Couturier, A., Maffiuletti, N.A., 2014. Validity of trunk extensor and flexor torque measurements using isokinetic dynamometry. J. Electromyogr. Kinesiol. 24 (6), 986–993.

Haff, G.G., 2018. 10 Strength–Isometric and Dynamic Testing. Performance Assessment in Strength and Conditioning.

Harding, A.T., Weeks, B.K., Horan, S.A., Little, A., Watson, S.L., Beck, B.R., 2017. Validity and test–retest reliability of a novel simple back extensor muscle strength test. SAGE Open Med. 5, 2050312116688842.

Hong, J., Reed, C., Novick, D., Happich, M., 2013. Costs associated with treatment of chronic low back pain: an analysis of the UK General Practice Research Database. Spine 38 (1), 75–82.

Hoy, D., Brooks, P., Blyth, F., Buchbinder, R., 2010. The Epidemiology of low back pain. Best Pract. Res. Clin. Rheumatol. 24 (6), 769–781.

Inc, B.M.S., 2020. Dual Position Back Ex/Flex Attachment Operation Manual. Biodex Medical Systems, Inc Shirley, NY, USA.

Jenkinson, C., Wright, L., Coulter, A., 1994. Criterion validity and reliability of the SF-36 in a population sample. Qual. Life Res. 3 (1), 7–12.

Johnson, K.D., Kim, K.-M., Yu, B.-K., Saliba, S.A., Grindstaff, T.L., 2012. Reliability of thoracic spine rotation range-of-motion measurements in healthy adults. J. Athl. Train. 47 (1), 52–60.

Jubany, J., Busquets, A., Marina, M., Cos, F., Angulo-Barroso, R., 2015. Reliability and validity of a custom-made instrument including a hand-held dynamometer for measuring trunk muscle strength. J. Back Musculoskelet. Rehabil. 28 (2), 317–326.

Karthikbabu, S., Chakrapani, M., 2017. Hand-held dynamometer is a reliable tool to measure trunk muscle strength in chronic stroke. J. Clin. Diagn. Res.: J. Clin. Diagn. Res. 11 (9), YC09.

Kato, S., Murakami, H., Demura, S., Yoshioka, K., Shinmura, K., Yokogawa, N., et al., 2019. Abdominal trunk muscle weakness and its association with chronic low back pain and risk of falling in older women. BMC Muscoskel. Disord. 20 (1), 1–8.

Kienbacher, T., Kollmitzer, J., Anders, P., Habenicht, R., Starek, C., Wolf, M., et al., 2016. Age-related test-retest reliability of isometric trunk torque measurements in patiens with chronic low back pain. J. Rehabil. Med. 48 (10), 893–902.

Kocjan, A., Sarabon, N., 2014. Assessment of isometric trunk strength–the relevance of body position and relationship between planes of movement. J. Sports Sci. Med. 13 (2), 365.

Konrad, P., 2005. The abc of emg. Pract. Intro. Kinesiologic. Electromyography. 1, 30–35, 2005.

Kroemer, K., Marras, W.S., 1980. Towards an objective assessment of the "maximal voluntary contraction" component in routine muscle strength measurements. Eur. J. Appl. Physiol. Occup. Physiol. 45 (1), 1–9.

Kumar, S., 1997. Axial rotation strength in seated neutral and prerotated postures of young adults. Spine 22 (19), 2213–2221.

Langrana, N.A., Lee, C.K., Alexander, H., Mayott, C., 1984. Quantitative assessment of back strength using isokinetic testing. Spine 9 (3), 287–290.

Mavroidis, C., Nikitczuk, J., Weinberg, B., Arango, R., Danaher, G., Jensen, K., et al. (Eds.), 2005. Smart Portable Rehabilitation Devices. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.

Menezes Costa, LdC., Maher, C.G., Hancock, M.J., McAuley, J.H., Herbert, R.D., Costa, L. O., 2012. The prognosis of acute and persistent low-back pain: a meta-analysis. Can. Med. Assoc. J. 184 (11), E613–E624.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., et al., 2010. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J. Clin. Epidemiol. 63 (7), 737–745.

Moreland, J., Finch, E., Stratford, P., Balsor, B., Gill, C., 1997. Interrater reliability of six tests of trunk muscle function and endurance. J. Orthop. Sports Phys. Ther. 26 (4), 200–208.

Nuzzo, J.L., Taylor, J.L., Gandevia, S.C., 2019. CORP: measurement of upper and lower limb muscle strength and voluntary activation. J. Appl. Physiol. 126 (3), 513–543.

Ozcan Kahraman, B., Salik Sengul, Y., Kahraman, T., Kalemci, O., 2016. Developing a reliable core stability assessment battery for patients with nonspecific low back pain. Spine 41 (14), E844–E850.

Park, H-w, Baek, S., Kim, H.Y., Park, J.-G., Kang, E.K., 2017. Reliability and validity of a new method for isometric back extensor strength evaluation using a hand-held dynamometer. Annals Rehabil. Med. 41 (5), 793–800.

Pekünlü, E., Özsu, İ., 2014. Avoiding systematic errors in isometric squat-related studies without pre-familiarization by using sufficient numbers of trials. J. Hum. Kinet. 42, 201.

Pengel, L.H., Herbert, R.D., Maher, C.G., Refshauge, K.M., 2003. Acute low back pain: systematic review of its prognosis. Br. Med. J. 327 (7410), 323.

Sanderson, A., Martinez-Valdes, E., Heneghan, N.R., Murillo, C., Rushton, A., Falla, D., 2019. Variation in the spatial distribution of erector spinae activity during a lumbar endurance task in people with low back pain. J. Anat. 234 (4), 532–542.

Stark, T., Walker, B., Phillips, J.K., Fejer, R., Beck, R., 2011. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. PM&R. 3 (5), 472–479.

Steele, J., Bruce-Low, S., Smith, D., 2014. A reappraisal of the deconditioning hypothesis in low back pain: review of evidence from a triumvirate of research methods on specific lumbar extensor deconditioning. Curr. Med. Res. Opin. 30 (5), 865–911.

Steele, J., Fisher, J., Perrin, C., Conway, R., Bruce-Low, S., Smith, D., 2019. Does change in isolated lumbar extensor muscle function correlate with good clinical outcome? A secondary analysis of data on change in isolated lumbar extension strength, pain, and disability in chronic low back pain. Disabil. Rehabil. 41 (11), 1287–1295.

Swinkels-Meewisse, E., Swinkels, R., Verbeek, A., Vlaeyen, J., Oostendorp, R., 2003. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man. Ther. 8 (1), 29–36.
Tarca, B.D., Wycherley, T.P., Meade, A., Bennett, P., Ferrar, K.E., 2020. Validity and reliability of hand-held dynamometry for abdominal flexion muscular assessment. J. Sport Rehabil. 30 (2), 343–346.
Taylor, J.B., Goode, A.P., George, S.Z., Cook, C.E., 2014. Incidence and risk factors for first-time incident low back pain: a systematic review and meta-analysis. Spine J. 14 (10), 2299–2319.
Trevethan, R., 2017. Intraclass Correlation Coefficients: Clearing the Air, Extending Some Cautions, and Making Some Requests, 17. Health Services and Outcomes Research Methodology.
Valentin, G., Maribo, T., 2014. Hand-held dynamometry fixated with a tripod is reliable for assessment of back extensor strength in women with osteoporosis. Osteoporosis Int. 25 (8), 2143–2149.
Verbrugghe, J., Agten, A., Eijnde, B.O., Vandenabeele, F., De Baets, L., Huybrechts, X., et al., 2019. Reliability and agreement of isometric functional trunk and isolated lumbar strength assessment in healthy persons and persons with chronic nonspecific low back pain. Phys. Ther. Sport 38, 1–7.
Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., et al., 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann. Intern. Med. 147 (8), 573–577.
Walter, S., Eliasziw, M., Donner, A., 1998. Sample size and optimal designs for reliability studies. Stat. Med. 17 (1), 101–110.
Yang, S., Wu, W., Zhang, C., Wang, D., Chen, C., Tang, Y., et al., 2020. Reliability and validity of three isometric back extensor strength assessments with different test postures. J. Int. Med. Res. 48 (2), 0300060519885268.