Physical Therapy in Sport 38 (2019) 1–7

Contents lists available at ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Original Research

Reliability and agreement of isometric functional trunk and isolated lumbar strength assessment in healthy persons and persons with chronic nonspecific low back pain

Jonas Verbrugghe [a, *], Anouk Agten [a], Bert O. Eijnde [a], Frank Vandenabeele [a], Liesbet De Baets [a], Xavier Huybrechts [b], Annick Timmermans [a]

[a] Hasselt University, REVAL - Rehabilitation Research Center, BIOMED, Hasselt, Belgium
[b] Department of Rehabilitation and Physical Medicine, Jessa Ziekenhuis, Hasselt, Belgium

ARTICLE INFO

Article history:
Received 11 January 2019
Received in revised form
20 March 2019
Accepted 20 March 2019

Keywords:
Chronic low back pain
Muscle strength
Reliability
Isokinetic dynamometer

ABSTRACT

Objectives: to assess intra/inter-operator reliability and agreement of maximum isometric abdominal and back muscle strength in a functional trunk and isolated lumbar protocol, using an isokinetic dynamometer, in healthy persons and persons with chronic nonspecific low back pain (CNSLBP).
Design: Test-retest.
Setting: Participants performed two assessments consisting of two protocols on the Biodex 3 system, evaluating maximum isometric back and abdominal strength in a functional trunk and isolated lumbar position. During the first assessment, each protocol was executed twice, supported by different operators.
Participants: Healthy persons (n = 20) and persons with CNSLBP (n = 20).
Main outcome measures: Intraclass Correlation (ICC), Standard Error of Measurement (SEM and %SEM), and Minimal Detectable Change (MDC) of muscle strength outcomes and seat positioning characteristics were calculated.
Results: Intra/inter reliability of muscle strength outcomes was excellent (ICC: 0.94–0.98), while seat positioning characteristics varied from low to high (ICC: 0–0.94). For muscle strength outcomes, %SEM ranged from 4.7 to 9.2% and MDC ranged from 14.3 to 29.8 Nm in trunk flexion and 39.1–68.5 Nm in trunk extension.
Conclusions: The Biodex 3 system can be used reliably to assess maximum isometric trunk muscle strength with the aforementioned protocols in healthy persons and persons with CNSLBP. All muscle strength outcomes showed comparable agreement (%SEM < 10%).

1. Introduction

Chronic low back pain is one of the most common musculoskeletal disorders worldwide (Hoy et al., 2012), occurring in persons of all age groups with a peak incidence between 30 and 65 years (Hoy et al., 2010). It can lead to high levels of functional disability and psychologic distress (Manchikanti et al., 2014) and gives rise to a high economic burden on healthcare systems (Dagenais, Caro, & Haldeman, 2008). Whereas 10–15% of persons with chronic low back pain can be diagnosed with a specific underlying pathology (e.g. disc herniation, canal stenosis, spondylolysis), most persons are categorized as having chronic nonspecific low back pain (CNSLBP), indicating that symptoms are of an unclear origin (Airaksinen et al., 2006).

In persons with CNSLBP, back and abdominal muscle functioning can be altered as decreased endurance and strength of trunk muscles have been noted (Behennah et al., 2018; Kankaanpää et al., 1998; Lee, Ooi, & Nakamura, 1995; Steele, Bruce-Low, & Smith, 2014). However, proper trunk muscle functioning is needed to provide trunk stabilization and movement control, and to minimize forces on the spine (Bergmark, 1989; Ebenbichler et al., 2001). Although the function of trunk strength as a risk factor for CNSLBP is debatable (Cho et al., 2014; McGill, 2015), reduced trunk extensor strength has been shown to be associated with the severity and

* Corresponding author. Agoralaan 5, Gebouw A 3590, Diepenbeek, Belgium.
E-mail address: Jonas.verbrugghe@uhasselt.be (J. Verbrugghe).

https://doi.org/10.1016/j.ptsp.2019.03.009
1466-853X/

level of disability in low back pain (Bayramoglu et al., 2001; Cho et al., 2014; Iwai et al., 2004). Moreover, rehabilitation aimed at strengthening of trunk region positively affects pain intensity and disability in persons with CNSLBP(Helmhout et al., 2017; Steele et al., 2018).

Thus, objective muscle strength assessment can be valuable to evaluate potential trunk muscle weakness and to support the planning of a rehabilitation process (Harbo, Brincks, & Andersen, 2012; Maud & Foster, 2006). Clinical, easy-to-perform-and-assess maximum effort tests have been developed to assess the trunk muscle performance in persons with CNSLBP (e.g. a static Sorensen test or a dynamic Arch up test). But, data on their validity is conflicting (Demoulin et al., 2012; Villafañe et al., 2016). Furthermore, they mainly evaluate muscle endurance and are not adapted to assess specific muscle strength (Demoulin et al., 2012). Isokinetic dynamometers or 'iso-machines' have been developed to provide a controlled, standardized, safe, and quantifiable solution for muscle strength assessment through digital force or torque tracking on an axis of rotation (Demoulin et al., 2012; Maud & Foster, 2006). Various iso-machines with adjustable designs are available to evaluate trunk strength and specific factors influencing it (e.g. differences in posture, hip angle, body stabilization) (Demoulin et al., 2012). Both static (isometric) and dynamic (isokinetic) strength tests can be executed, although standard methods for isokinetic testing are not yet agreed upon (Demoulin et al., 2012).

Due to the specificity of these iso-machines and their protocols, identification of the psychometric properties is increasingly important. As such, the Biodex 3 system provides a distinct attachment for muscle strength evaluation of back extension and abdominal flexion (Biodex Medical Systems). Gabiner et al.(Grabiner, Jeziorowski, & Divekar, 1990) already stated guidelines on how to optimize the seating position of this system for reliability in isometric and isokinetic trunk flexion and extension testing and Van Dieen & Heijblom (van Dieën & Heijblom, 1996) showed reproducibility (test-retest reliability) of maximum isometric back extension in healthy persons. The capability of this attachment to gain specific information in both a functional trunk and isolated lumbar position holds an added advantage, as Steele et al. noted the importance of a clear differentiation between trunk extension ('TEX', i.e. a position allowing rotation of the pelvis and increased influence of hip extensor musculature) and isolated lumbar extension ('ILEX', i.e. a position with pelvic stabilization limiting accessory muscle work) when performing lumbar spine strength assessment (Steele et al., 2014). However, in all previous studies, protocol specifics of the testing procedure were unclear and no differentiation was made concerning reliability of the specific TEX and ILEX positions. Furthermore, a study by Gruther et al.(Gruther et al., 2009) showed lower test-retest reliability in the evaluation of persons with CNSLBP although it suffered from the same limitations as mentioned above. Besides, no agreement outcomes are currently available with regard to maximum back and abdominal muscle strength assessment in healthy persons and persons with CNSLBP, making it difficult to evaluate the outcomes of trunk strength tests in a rehabilitation setting.

Therefore, the objective of this study is to assess intra/inter-operator reliability and agreement of maximum isometric muscle strength of the back and abdominal muscles in a functional trunk and isolated lumbar protocol, using the Biodex system 3 isokinetic dynamometer, in healthy persons and persons with CNSLBP.

2. Methods

2.1. Participants

A formula proposed by Bujang & Baharum (Bujang & Baharum, 2017) was used to calculate the minimal sample size for healthy persons (HP) and persons with chronic nonspecific low back pain (CNSLBP). This formula showed that a minimal sample size of 15 was sufficient to detect values of ICC >0.6 (i.e. moderate correlation (Shrout, 1998)), with alpha (type 1 error) set at <0.05 and power (type 2 error) > 80% for a test-retest and intra-operator reliability design (where the agreement in the null hypothesis (R0) is pre-specified to be equal to 0.0). To account for a 10% drop-out rate, a minimum of 18 persons needed to be tested in each group (Bujang & Baharum, 2017). Therefore, twenty HP and twenty persons with CNSLBP were recruited.

HP were recruited via family and acquaintances of the researchers. To be eligible, HP had to be 25–60 years old, understand verbal instructions and be able to fill in questionnaires in Dutch. Chronic disorders influencing the ability to perform maximal strength testing of back and abdominal muscles, co-morbidities (e.g. paresis or sensory disturbances by neurological causes), and pregnancy were exclusion criteria.

Persons with CNSLBP were recruited by the researchers through local flyer distribution, and by physiotherapists of X Hospital (X). To be eligible, persons with CNSLBP had to be 25–60 years old, have chronic low back pain of a nonspecific origin (i.e. pain localized below the costal margin and above the inferior gluteal folds with or without referred leg pain of a nociceptive mechanical nature, not attributable to a recognizable, known specific pathology e.g. infection, tumor, osteoporosis, fracture, structural deformity, inflammatory disorder, radicular syndrome, or cauda equina syndrome, for a period of at least twelve weeks) (Airaksinen et al., 2006, Balagué, Mannion, Pellisé, & et al, 2012), understand verbal instructions and be able to fill in questionnaires in Dutch. Co-morbidities (e.g. paresis or sensory disturbances by neurological causes), a pain score of >8/10 in the last 24 h, or pregnancy, were exclusion criteria. Also, persons with CNSLBP were not allowed to perform any type of exercise therapy prior to or between the measurements.

Persons with CNSLBP who wanted to participate were invited for an intake session before the first assessment. At the start of this intake session, participants had to provide a medical diagnosis of their low back pain by a GP or specialist. Then the information letter was reviewed, study in- and exclusion criteria were evaluated, the informed consent was signed, and a study specific screening form concerning low back pain symptoms was filled out.

2.2. Study design

Each participant performed two separate maximum isometric trunk muscle strength assessments at X (X), five to ten days apart, supported by the same operator (JV). During the first assessment, the testing protocol was performed twice (after a resting period of 30 min to avoid muscle fatigue during the second protocol), supported by a second operator (AA). Both operators had extensive experience in performing muscle strength testings and using the instruments described below. To evaluate intra-operator reliability, outcomes of JV of the first assessment were compared with outcomes of JV of the second assessment. To evaluate inter-operator reliability, outcomes of operator JV were compared to outcomes of operator AA. This study was approved by the medical ethical committee of X (X) and X Hospital (X) and was registered at clinicaltrials.gov (X). All participants provided written informed consent.

2.3. Instruments and measurements

Maximum isometric back and abdominal muscle strength was measured using an isokinetic dynamometer (Biodex system 3 pro, Dual Position Back Extension/Flexion Attachment, Enraf-Nonius, USA) (Drouin et al., 2004). The participant was fixated in an adjustable seat in two standardized positions i.e. 'semiflex' and 'isolated lumbar' (Fig. 1B and C). These positions allowed for the execution of two protocols evaluating back (extension) and abdominal (flexion) muscle performance in either functional or isolated capacities (Biodex Medical Systems). The functional 'semiflex' protocol was set up to evaluate generalized trunk extension (TEX, i.e. a compound movement requiring additional rotation of the pelvis through the hip extensor musculature), while the 'isolated lumbar' protocol was set up to evaluate specific isolated lumbar extension (ILEX, i.e. an isolated movement requiring utilizing pelvic stabilization through use of a semi-seated position with rear pelvic restraint and a belt across the thighs) (Steele et al., 2014). Detailed protocol descriptions are provided in appendix A. For each protocol, following muscle strength outcomes were inventoried: maximum absolute isometric peak torque for extension ('EXT', in Nm) and flexion ('FLE', in Nm). For each protocol following seat position outcomes were inventoried: seat height ('height', in cm), scapula roll ('roll', in units), hip angle ('angle', in °) and sacral pad ('depth', in cm) (only in isolated lumbar). Table 1 provides the unit and range of each seat position outcome. Fig. 1A displays a visual representation of the seat position outcomes.

2.4. Procedures

Following baseline participant characteristics were collected: length (cm), weight (kg), and activity level (Physical activity Scale for Individuals with Disabilities). In persons with CNSLBP, pain (Numeric pain rating score) and disability (Modified Oswestry Disability Index) questionnaires were also filled out. Then, participants performed an 8 min warm-up of brisk walking on a treadmill (4-7 kph) (Maud & Foster, 2006; Roussel et al., 2006). Subsequently, participants were seated on the Biodex system. Table 1 provides a detailed description of each seat position, based on previous research (Grabiner et al., 1990). As a specific trunk warm-up and to familiarize participants with the system, fifteen repetitions of dynamic concentric trunk flexion and extension were performed at low intensity (Maud & Foster, 2006). Next, participants were educated on the correct execution of the maximum protocols. The semiflex protocol was first, followed by the isloated lumbar protocol. During both protocols, three repetitions of maximum isometric peak extension and three repetitions of maximum flexion torque were recorded (Hopkins, 2000; Maud & Foster, 2006). For each repetition, participants were provided with a 10 s time slot to gradually increase force production until achieving and holding a maximum peak torque value for three to 5 s (Cho et al., 2014; Maud & Foster, 2006). Extension and flexion repetitions were alternated. Between each repetition a resting period of 30 s was given for the participant to prepare for the next repetition. Prior to executing the maximum protocols, a sub-maximum simulation protocol of two repetitions of isometric trunk flexion and extension was performed, to limit learning effect during the following maximum protocols (Grabiner et al., 1990; van Dieën & Heijblom, 1996).

2.5. Statistical analysis

SPSS version 23 (Chicago, IL) was used to analyze data. Normal distribution of the data was checked (Shapiro-wilk test) and means and standard deviations were calculated. Intraclass Correlation Coefficients (ICCs) were calculated to evaluate inter- (ICC_b) and intra-operator (ICC_w) reliability, using a two-way mixed model with absolute agreement (ICC type 3,1) (Koo & Li, 2016) and 95% confidence intervals (95%CI). ICCs were interpreted according to) (Shrout, 1998): values > 0.80 represent substantial, 0.61–0.80 moderate, 0.41–0.60 fair, 0.11–0.40 slight and 0–0.10 virtually no reliability. The standard error of measurement (SEM) was calculated for inter- (SEM_b) and intra-operator (SEM_w) agreement as follows: $SEM = \sqrt{\text{mean square error (MSE)}}$ (Weir, 2005). Proportional SEM was calculated by expressing SEM relative to the mean. The minimal detectable change (MDC) between two sessions was calculated using $SEM^{*}1.96^{*}\sqrt{2}$ (Weir, 2005). Bland-Altmann plots were constructed to visualize the variability of the measurements by showing fluctuations around the mean. For each comparison, the 95% limit of agreement (LOA) was computed as follows: $LOA = \text{mean difference} \pm (1{,}96^{*}SD_{meandifference})$. A one sample t-tests or Wilcoxon signed ranked tests (for non-normally distributed data) using the inter- and intra-operator mean differences of the strength outcomes were performed to check for systematic bias in these datasets.

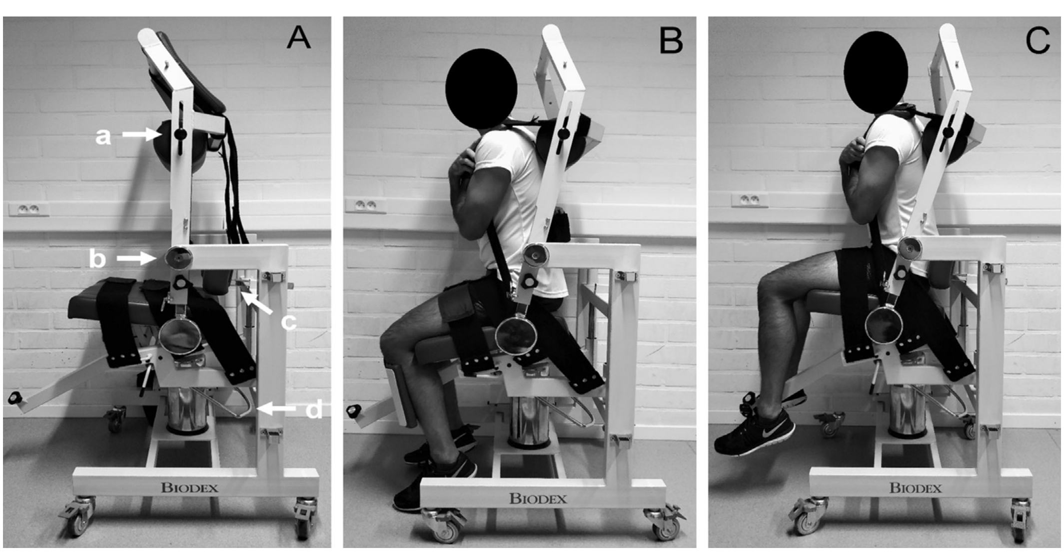

Fig. 1. A: seat position outcomes (a: scapula roll; b: angle; c: lumbar pad; d: seat height); B: semiflex protocol specific participant positioning; and C: isloated lumbar protocol specific participant positioning.

Table 1
Detailed description of the seating position outcomes.

Seat position	Outcome	Range	Description
Seat height ('height')	Centimeter (cm)	0.0–18.5	Seat height is adjusted to align the fixed axis of the system with the subject's anterior superior iliac spine (ASIS) on the sagittal axis.
Scapula roll ('roll')	Units	1–4	The upper border of the shoulder roll is aligned to rest between the level of the scapular spines and inferior angles.
Seat angle ('angle')	Degrees (°)	0–90	A hip angle of 90° is measured by using a manual goniometer from the horizontal plane of the seat perpendicular to the middle of the palpated acromion of the participant. The angle of the chair is then retrieved from the internal Biodex goniometer.
Lumbar pad ('depth')	Units	1–4	Seat depth is adjusted to align the fixed axis of the machine with the subject's anterior superior iliac spine (ASIS) on the vertical axis.

3. Results

Twenty HP (age = 41.1 y (12.3), 11 females) and twenty persons with CNSLBP (age = 44.0 y (11.2), 10 females) were tested. One HP did not complete the study due to acute low back pain during the execution of the first isolated lumbar protocol on assessment one. Therefore, the dataset of this participant was not used for further analysis. An overview of general patient characteristics is given in Table 2. An overview of Means (SD), ICCs (95% confidence intervals), SEMs, and MDCs of both HP and CNSLBP is reported in Table 3.

In healthy persons, muscle strength outcomes $ICCs_b$ and $ICCs_w$ showed substantial reliability (0.86–0.97) in both the semiflex and isolated lumbar protocol. SEMs ranged from 14.1 to 22.1Nm in EXT (%SEM 5.9–8.5) and 5.2 to 10.7Nm in FLE (%SEM 4.9–8.2). Corresponding MDCs ranged from 39.1 to 61.3Nm in EXT and 14.3 to 29.8Nm in FLE. Seat position outcomes $ICCs_w$ were fair to substantial (0.42–0.82), with the only exception of depth (0.10). $ICCs_b$ were fair to substantial (0.51–0.81) with the exception of angle and depth (0.23 and 0). Angle SEMs ranged from 2.8 to 4.0° (%SEM 4.0–7.8), corresponding with MDCs of 0.8–8.9°. Height SEMs ranged from 0.5 to 1.2 cm (%SEM 3.1–9.8), corresponding with MDCs of 1.4–3.4 cm. Roll SEMs ranged from 0 to 0.5 units (%SEM 0–13.9), corresponding with MDCs of 0–1.4 units. Depth SEM in isolated lumbar was 0.3 units (%SEM 9.1–9.3), corresponding with an MDC of 0.7–0.8 units.

In persons with nonspecific chronic low back pain, muscle strength outcome $ICCs_b$ and $ICCs_w$ showed substantial reliability (0.89–0.98) in both the semiflex and isolated lumbar protocol. SEMs ranged from 20.2 to 24.7Nm in EXT and 5.9 to 9.8Nm in FLE. Corresponding MDCs ranged from 56.1 to 68.5Nm in EXT and 16.4 to 27.0Nm in FLE. Seat position outcomes $ICCs_w$ were fair to moderate (0.34–0.76), while $ICCs_b$ were slight to substantial (0.16–0.81). Angle SEMs ranged from 2.8 to 4.0°, corresponding with MDCs of 0.8–8.9°. Height SEMs ranged from 0.3 to 1.3 cm, corresponding with MDCs of 0.9–3.7 cm. Roll SEMs ranged from 0.2 to 0.3 units, corresponding with MDCs of 0.6–0.8 units. Depth SEMs in isolated lumbar ranged from 0.1 to 0.3 units, corresponding with MDCs of 0.4–0.8 units.

Bland-Altman plots for intra/inter-operator agreement assessment are provided in Appendix B. No significant differences between mean differences were found for the inter- or intra-operator strength outcomes.

4. Discussion

This study aimed to evaluate intra/inter-operator reliability and agreement of isometric muscle strength assessment of trunk flexion and extension in a functional trunk and isolated lumbar protocol, using the Biodex system 3, in healthy persons (HP) and persons with chronic nonspecific low back pain (CNSLBP). Results showed substantial intra- and inter-operator ICCs in as well the semiflex as the isolated lumbar protocol for all muscle strength outcomes, in both HP and in persons with CNSLBP. This is in line with previous research showing reliability of isometric and isokinetic trunk muscle testing in other populations using this system (García-Vaquero et al., 2016; van Dieën & Heijblom, 1996), and reliability of other iso-machines (e.g. David, IsoMed 2000) to perform isometric trunk strength testing (Kienbacher et al., 2016; Roth et al., 2017). On the contrary, comparable studies evaluating iso-machine assisted maximum muscle strength with less reliable outcomes have also been noted (Madsen, 1996). This discrepancy in results is thought to be due to methodological limitations such as improper patient stabilization, axis placement, and unclear definition of testing terms (Newton & Waddell, 1993; Steele et al., 2014). In addition, most studies only display test-retest (intra-operator) reliability (Newton & Waddell, 1993). Therefore, the current study evaluated both inter- and intra-rater reliability and made rigorous efforts to improve reproducibility by providing transparent testing protocols and seat position characteristics. Also, as both Grabiner et al.(Grabiner et al., 1990) and van Dieen & Heijblom (van Dieën & Heijblom, 1996) discussed the influence of learning effect during isometric strength testing, this study used an extended warm up and simulation protocol to avoid learning effect in affecting the outcomes. Furthermore, this study evaluated two

Table 2
Patient characteristics.

	HP (n = 19)			CNSLBP (n = 20)		
Outcome	Mean	SD	Range	Mean	SD	Range
Gender (M/F)	8/11	–	–	10/10	–	–
Age (Y)	40.2	11.9	24–57	44,0	11.2	21–59
Length (*Cm*)	171.3	9.8	157.2–190.0	176,2	8,4	161.5–189.0
Weight (Kg)	67.5	10.7	53.88.0	79,3	15,8	52.4–111.8
Activity Level (PASIPD)	20.4	10.9	2.5–41.4	17,9	10,8	2.2–34.5
Pain Intensity (NPRS)				5,1	1,8	2.0–8.0
Disability (MODI)				11,5	6,3	2–24
Kinesiophobia (TSK)				34,5	7,2	23–46

Abbreviations: HP: healthy persons; CNSLBP: chronic nonspecific low back pain; PASIPD: physical activity scale for individuals with disabilities; NPRS: numeric pain rating scale; MODI: modified oswestry disability index; TSK: Tampa Scale for Kinesiophobia.

Table 3
Intra/inter-operator means, standard deviations, intraclass correlations, 95% confidence intervals, standard error of measurements and minimal detectable changes for both healthy persons (HP) and persons with chronic nonspecific low back pain (CNSLBP).

HP (n = 19)															
Intra-operator								Inter-operator							
Outcome	Mean	SD	ICC_w	95%CI	SEM	%SEM	MDC_w	Outcome	Mean	SD	ICC_b	95%CI	SEM_b	%SEM	MDC
semiflex															
EXT	238.3	89.3	0.94	(0.85, 0.98)	22.1	8.9	61.3	EXT	247.8	92.8	0.95	(0.87, 0.98)	21.0	8.5	58.0
FLE	150.6	55.5	0.97	(0.91, 0.99)	10.7	5.8	29.8	FLE	152.9	55.9	0.97	(0.91, 0.99)	9.2	6.0	25.6
Angle	66.7	3.6	0.42	(0.00, 0.73)	2.7	4.0	7.5	Angle	67.2	4.3	0.23	(0.00, 0.62)	3.8	5.7	10.6
Height	12.0	1.2	0.62	(0.24, 0.84)	0.7	5.8	2.1	Height	12.2	1.7	0.51	(0.10, 0.78)	1.2	9.8	3.4
Rol	3.7	0.6	0.94	(0.84, 0.97)	0.5	13.5	1.4	Rol	3.6	0.7	0.95	(0.84, 0.97)	0.5	13.9	1.4
Isolated lumbar															
EXT	232.2	86.7	0.97	(0.93, 0.99)	14.8	6.4	40.9	EXT	239.0	87.7	0.97	(0.92, 0.99)	14.1	5.9	39.1
FLE	104.2	35.4	0.94	(0.84, 0.98)	8.5	8.2	23.4	FLE	106.3	34.1	0.94	(0.35, 0.99)	5.2	4.9	14.3
Angle	50.1	5.5	0.51	(0.09, 0.78)	3.9	7.8	10.9	Angle	50.5	4.2	0.70	(0.39, 0.87)	2.2	4.4	6.2
Height	16.2	1.2	0.82	(0.60, 0.93)	0.5	3.1	1.5	Height	16.2	1.3	0.80	(0.56, 0.92)	0.6	3.7	1.6
Rol	3.7	0.6	1	–	–	–	–	Rol	3.7	0.6	0.92	(0.81, 0.97)	0.2	5.4	0.5
Depth	3.2	0.3	0.10	(0.00, 0.52)	0.3	9.4	0.7	Depth	3.3	0.3	0	(0.00, 0.31)	0.3	9.1	0.8
Persons with CNSLBP (n = 20)															
Intra-operator								Inter-operator							
Outcome	mean	SD	ICC	95%CI	SEM	%SEM	MDC	outcome	mean	SD	ICC	95%CI	SEM	%SEM	MDC
Semiflex															
EXT	269.5	95.6	0.94	(0.86, 0.98)	24.1	8.9	66.9	EXT	268.3	93.1	0.94	(0.84, 0.97)	24,7	9.2	68,5
FLE	155.2	58.3	0.98	(0.95, 0.99)	9.0	6.0	25.0	FLE	155.4	55.4	0.97	(0.93, 0.99)	9,8	6.3	27,0
Angle	66.6	3.7	0.44	(0.00, 0.73)	2.8	4.2	7.8	Angle	65.2	4.3	0.36	(0.00, 0.67)	3,2	4.9	8,9
Height	12.6	1.3	0.75	(0.46, 0.89)	0.7	5.6	1.9	Height	12.5	1.4	0.16	(0.00, 0.56)	1,3	10.4	3,7
Rol	3.4	0.5	0.75	(0.46, 0.89)	0.3	8.8	0.8	Rol	3.4	0.6	0.79	(0.53, 0.91)	0,3	8.8	0,8
Isolated lumbar															
EXT	244.0	83.3	0.93	(0.83, 0.97)	22.6	9.3	62.7	EXT	249.2	87.3	0.97	(0.87, 0.98)	20.2	8.1	56,1
FLE	108.2	32.1	0.97	(0.92, 0.99)	5.9	5.5	16.4	FLE	108.4	34.5	0.94	(0.85, 0.98)	9.1	8.4	25,2
Angle	49.7	4.8	0.34	(0.00, 0.68)	4.0	8.0	11.0	Angle	50.8	4.3	0.44	(0.01, 0.74)	2,77	5.5	7,7
Height	16.7	1.2	0.91	(0.77, 0.97)	0.3	1.8	0.9	Height	16.5	1.1	0.65	(0.31, 0.84)	0,66	4.0	1,8
Rol	3.5	0.5	0.81	(0.59, 0.92)	0.2	5.7	0.6	Rol	3.6	0.5	0.81	(0.57, 0.92)	0,23	6.4	0,6
Depth	3.1	0.3	0.76	(0.48, 0.90)	0.1	3.2	0.4	Depth	3.1	0.4	0.31	(0.00, 0.65)	0,30	9.7	0,8

Abbreviations: ICC: Intraclass Correlation Coefficient; CI: confidence interval; SEM: standard error of measurement; MDC: minimal detectable change; AWY: isometric back strength in Newtonmeters (nm); FWD: isometric abdominal strength in newtonmeters (Nm); angle: hip angle in degrees (°); height: seat height in centimeters (cm); rol: scapula rol in units from 1 to 4; depth: lumbar pad in units from 1 to 4.

differentiating protocols, that could provide more nuanced information concerning muscle strength in these populations (Biodex Medical Systems). As the semiflex protocol is displayed as a functional protocol (i.e. positioning with lumbar and sacral movement and contribution of posture musculature in trunk extension, and hip muscle involvement in trunk flexion (Biodex Medical Systems)), higher outcomes were expected in this protocol compared to the isolated lumbar protocol ('ILEX', i.e. positioning with pelvic stabilization limiting accessory muscle work in both lumbar extension and flexion). However, when comparing both protocols, differences between muscle strength outcomes in trunk extension measured in the semiflex and isolated lumbar protocol were small. Higher outcomes were only found in the semiflex protocol for functional trunk flexion. The authors recommend future research to validate these protocols and further evaluate differences in strength between protocols, as this was beyond the scope of the current study.

Agreement was evaluated by means of standard error of measurement (SEM) and minimal detectable change (MDC) (Weir, 2005). SEM shows the influence of possible subject homogeneity and between-subject variability on the outcomes of the ICCs. MDC has clinical relevance as it can reflect true change in an outcome (i.e. smallest amount of change that can be considered above the threshold of error expected in the measurement) (Portney & Watkins, 2009). Both SEM and MDC provide an absolute measure of reliability instead of a proportional, by using the same units as the test of interest. Thus, they are more applicable when using the proposed protocols in a clinical setting, as they provide a direct evaluation reference point and can showcase progression when patients are assessed repeatedly. All muscle strength outcomes in both protocols showed comparable inter- and intra-operator agreement (%SEM < 10%) in HP and persons with CNSLBP. Therefore, these protocols show usability in the evaluation of physical fitness or rehabilitation programs, as in untrained to moderately trained individuals strength increases of 20–40% can be expected by executing a resistance training program (Maud & Foster, 2006). In persons with comparable characteristics as this sample, CNSLBP as a pathology does not seem to influence the reliability of maximum isometric trunk strength testing. However, the mean pain intensity and disability level of the persons with CNSLBP in this study were only low to moderate. In a population that is more pain or disability dominant, influence of these outcomes could differ. Also, other specific outcomes such as a high levels of kinesiophobia, maladaptive coping, and central sensitization have also been noted in CNSLP (Ramond et al., 2011; Roussel et al., 2013). Future research should further evaluate the influence of these outcomes on trunk muscle strength testing.

Aside from these results, the authors note some methodological considerations. Firstly, some of the seat position outcomes showed average to low reliability (least reliable were 'angle' and 'depth'). Possibly, this was due to the fact that these outcomes were collected manually by each operator (e.g. use of a manual goniometer to measure the angle), opposed to the muscle strength outcomes which were directly recorded by the system itself. However, these outcomes were administered to account for possible variability issues in muscle strength outcomes. As muscle strength outcomes were all substantially reliable and differences in mean of seat position outcomes were clinically very small (e.g. max. 1.4° in angle, 0.2 cm in height), the authors conclude that although correct initial patient positioning remains important, differences in seat position between or within raters will only have a minor impact when evaluating maximum isometric strength with these protocols. Secondly, pain intensity in persons with CNSLBP was only administered at the start of the first assessment. Thus, a change in pain intensity during the assessment or on the day of the second testing moment could have affected the strength outcomes. However, no person with CNSLBP complained of increased pain during or after the protocols or aborted any of the assessments due to pain. Besides, no significant differences in strength were measured between both testing moments. This leads the authors to assume that pain did not affect the outcomes. Thirdly, as males show a higher performance on maximum strength measures than females (Newton & Waddell, 1993), primarily due to a difference in mean lean mass affecting the ability to generate force (Chen et al., 2013), normalization for body weight has been executed in a number of studies. Thus, the authors of this study have performed a post hoc analysis for relative muscle strength outcomes (Nm/kg) on the dataset of this study, to evaluate the influence of weight on reliability outcomes (data not shown). Intra- and inter-operator reliability for all muscle strength outcomes were almost identical, and also SEM and MDC were highly comparable. Therefore, differences in weight do not seem to affect reliability when using the aforementioned protocols. Lastly, as the scapula roll was situated between the lower end of the spine of the scapula and the inferior angle of the scapula, the depicted back muscle strength could be partly due to back muscles from the thoracic region. An evaluation that provides fixation specifically between vertebral levels T12 and S1 or a research evaluating muscle activity during isometric muscle strength assessment (e.g. concurrent EMG analysis) could provide more knowledge concerning this limitation.

5. Conclusion

This study showed substantial intra- and inter-operator reliability of maximum isometric back extension and abdominal flexion in a functional trunk and isolated lumbar protocol, in both healthy persons and persons with chronic nonspecific low back pain by using the Biodex system 3. All muscle strength outcomes showed comparable inter- and intra-operator agreement (% SEM < 10%) in both HP and persons with CNSLBP.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ptsp.2019.03.009.

References

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. *European Spine Journal, 15*, S192–S300.

Balagué, F., Mannion, A. F., Pellisé, F., & Cedraschi, C. (2012). Non-specific low back pain. *Lancet, 379*, 482–491.

Bayramoglu, M., Akman, M. N., Klnç, S., Çetin, N., Yavuz, N., & Özker, R. (2001). Isokinetic measurement of trunk muscle strength in women with chronic low-back pain. *American Journal of Physical Medicine & Rehabilitation, 80*, 650–655.

Behennah, J., Conway, R., Fisher, J., Osborne, N., Steele, J., et al. (2018). The relationship between balance performance, lumbar extension strength, trunk extension endurance, and pain in participants with chronic low back pain, and those without. *Clinical Biomechanics, 53*, 22–30.

Bergmark, A. (1989). Stability of the lumbar spine: A study in mechanical engineering. *Acta Orthopaedica Scandinavica, 60*, 1–54.

Biodex Medical Systems, I.. Dual position back ex/flex attachment operation manual. Available on: http://m.biodex.com/sites/default/files/830450man_08261reva.pdf.

Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Sciences, 12*.

Chen, L., Nelson, D. R., Zhao, Y., Cui, Z., & Johnston, J. A. (2013). Relationship between muscle mass and muscle strength, and the impact of comorbidities: A population-based, cross-sectional study of older adults in the United States. *BMC Geriatrics, 13*, 74.

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. *Annals of Rehabilitation Medicine, 38*, 234–240.

Dagenais, S., Caro, J., & Haldeman, S. (2008). A systematic review of low back pain cost of illness studies in the United States and internationally. *The Spine Journal, 8*, 8–20.

Demoulin, C., Grosdent, S., Smeets, R., Verbunt, J., Jidovtseff, B., Mahieu, G., et al. (2012). *Muscular performance assessment of trunk extensors: A critical appraisal of the literature in low back pain*. InTech.

van Dieën, J. H., & Heijblom, P. (1996). Reproducibility of isometric trunk extension torque, trunk extensor endurance, and related electromyographic parameters in the context of their clinical applicability. *Journal of Orthopaedic Research, 14*, 139–143.

Drouin, J. M., Valovich-mcLeod, T. C., Shultz, S. J., Gansneder, B. M., & Perrin, D. H. (2004). Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. *European Journal of Applied Physiology, 91*, 22–29.

Ebenbichler, G. R., Oddsson, L. I., Kollmitzer, J., & Erim, Z. (2001). Sensory-motor control of the lower back: Implications for rehabilitation. *Medicine & Science in Sports & Exercise, 33*, 1889–1898.

García-Vaquero, M. P., Barbado, D., Juan-Recio, C., López-Valenciano, A., & Vera-Garcia, F. J. (2016). Isokinetic trunk flexion–extension protocol to assess trunk muscle strength and endurance: Reliability, learning effect, and sex differences. *Journal of Sport and Health Science*.

Grabiner, M. D., Jeziorowski, J. J., & Divekar, A. D. (1990). Isokinetic measurements of trunk extension and flexion performance collected with the biodex clinical data station. *Journal of Orthopaedic & Sports Physical Therapy, 11*, 590–598.

Gruther, W., Wick, F., Paul, B., Leitner, C., Posch, M., Matzner, M., et al. (2009). Diagnostic accuracy and reliability of muscle strength and endurance measurements in patients with chronic low back pain. *Journal of Rehabilitation Medicine, 41*, 613–619.

Harbo, T., Brincks, J., & Andersen, H. (2012). Maximal isokinetic and isometric muscle strength of major muscle groups related to age, body mass, height, and sex in 178 healthy subjects. *European Journal of Applied Physiology, 112*, 267–275.

Helmhout, P. H., Witjes, M., Nijhuis-Van, R. D. S., Bron, C., & Staal, J. (2017). The effects of lumbar extensor strength on disability and mobility in patients with persistent low back pain. *The Journal of Sports Medicine and Physical Fitness, 57*, 411–417.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*, 1–15.

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., et al. (2012). A systematic review of the global prevalence of low back pain. *Arthritis & Rheumatism, 64*, 2028–2037.

Hoy, D., Brooks, P., Blyth, F., & Buchbinder, R. (2010). The epidemiology of low back pain. *Clinical Rheumatology, 24*, 769–781.

Iwai, K., Nakazato, K., Irie, K., Fujimoto, H., & Nakajima, H. (2004). Trunk muscle strength and disability level of low back pain in collegiate wrestlers. *Medicine & Science in Sports & Exercise, 36*, 1296–1300.

Kankaanpää, M., Taimela, S., Laaksonen, D., Hänninen, O., & Airaksinen, O. (1998). Back and hip extensor fatigability in chronic low back pain patients and controls. *Archives of Physical Medicine and Rehabilitation, 79*, 412–417.

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 patients with chronic low back pain. *Journal of Rehabilitation Medicine, 48*, 893–902.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Chiropractic Journal of Medicine, 15*, 155–163.

Lee, J.-H., Ooi, Y., & Nakamura, K. (1995). Measurement of muscle strength of the trunk and the lower extremities in subjects with history of low back pain. *Spine, 20*, 1994–1996.

Madsen, O. R. (1996). Trunk extensor and flexor strength measured by the cybex 6000 dynamometer: Assessment of short-term and long-term reproducibility of several strength variables. *Spine, 21*, 2770–2776.

Manchikanti, L., Singh, V., Falco, F. J., Benyamin, R. M., & Hirsch, J. A. (2014). Epidemiology of low back pain in adults. *Neuromodulation, 17*(Suppl 2), 3–10.

Maud, P. J., & Foster, C. (2006). Physiological assessment of human fitness. *Human Kinetics*.

McGill, S. M. (2015). *Low back disorders, 3E*. Human Kinetics.

Newton, M., & Waddell, G. (1993). Trunk strength testing with iso-machines. Part 1: Review of a decade of scientific evidence. *Spine, 18*, 801–811.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice*.

Ramond, A., Bouton, C., Richard, I., Roquelaure, Y., Baufreton, C., Legrand, E., et al. (2011). Psychosocial risk factors for chronic low back pain in primary care—a systematic review. *Family Practice, 28*, 12–21.

Roth, R., Donath, L., Kurz, E., Zahner, L., & Faude, O. (2017). Absolute and relative reliability of isokinetic and isometric trunk strength testing using the IsoMed-2000 dynamometer. *Physical Therapy in Sport, 24*, 26–31.

Roussel, N., Nijs, J., Meeus, M., Mylius, V., Fayt, C., & Oostendorp, R. (2013). Central sensitization and altered central pain processing in chronic low back pain: Fact or myth? *The Clinical Journal of Pain, 29*, 625–638.

Roussel, N., Nijs, J., Truijen, S., Breugelmans, S., Claes, I., & Stassijns, G. (2006). Reliability of the assessment of lumbar range of motion and maximal isometric strength. *Archives of Physical Medicine and Rehabilitation, 87*, 576–582.

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical*

Methods in Medical Research, 7, 301–317.

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. *Current Medical Research and Opinion, 30*, 865–911.

Steele, J., Fisher, J., Perrin, C., Conway, R., Bruce-Low, S., & Smith, D. (2018). 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. *Disability & Rehabilitation, 1–9*.

Villafañe, J. H., Gobbo, M., Peranzoni, M., Naik, G., Imperio, G., Cleland, J. A., et al. (2016). Validity and everyday clinical applicability of lumbar muscle fatigue assessment methods in patients with chronic non-specific low back pain: A systematic review. *Disability & Rehabilitation, 38*, 1859–1871.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research, 19*, 231–240.