Title

The reliability and validity of a designed setup for the assessment of static back extensor force and endurance in older women with and without hyperkyphosis

Permalink
https://escholarship.org/uc/item/1n92d66v

Journal
Physiotherapy Theory and Practice, 34(11)

ISSN
0959-3985

Authors
Roghani, Taybeh
Zavieh, Minoo Khalkhali
Rahimi, Abbas
et al.

Publication Date
2018-11-02

DOI
10.1080/09593985.2018.1430878

Peer reviewed

Published in final edited form as:
Physiother Theory Pract. 2018 November ; 34(11): 882–893. doi:10.1080/09593985.2018.1430878.

The reliability of a designed setup for the assessment of static back extensor force and endurance in older women with and without hyperkyphosis

Taybeh Roghani, PT, PhD[1] **[Candidate]**, **Minoo Khalkhali Zavieh, PT, PhD**[1], **Abbas Rahimi, PT, PhD**[1], **Saeed Talebian, PT, PhD**[1], **Farideh Dehghan Manshadi, PT, PhD**[1], **Alireza Akbarzadeh Baghban, PhD**[2], **Nicole King, PT**[3], and **Wendy Katzman, PT, DPTSc, OCS**[3]

[1]Department of Physiotherapy, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[2]Proteomics Research Center, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[3]Department of Physical Therapy and Rehabilitation Sciences, University of California, San Francisco, United Sates

Abstract

Objective: The purpose of this study was to investigate the intra-rater reliability and validity of a designed load cell setup for the measurement of back extensor muscle force and endurance.

Participants: The study sample included 19 older women with hyperkyphosis, mean age 67.0 ± 5.0 years, and 14 older women without hyperkyphosis mean age 63.0 ± 6.0 years.

Methods: Maximum back extensor force and endurance were measured in a sitting position with a designed load cell setup. Tests were performed by the same examiner on 2 separate days within a 72-hour interval. The intra-rater reliability of the measurements was analyzed using Intraclass correlation coefficient (ICC), standard errors of measurement (SEM) and minimal detectable change (MDC). The validity of the setup was determined using Pearson correlation analysis and independent t-test.

Results: Using our designed load cell, the values of ICC indicated very high reliability of force measurement (hyperkyphosis group: 0.96, normal group: 0.97) and high reliability of endurance measurement (hyperkyphosis group: 0.82, normal group: 0.89). For all tests, the values of SEM and MDC were low in both groups. A significant correlation between two documented forces (load cell force and target force) and significant differences in the muscle force and endurance among the two groups were found.

Conclusion: The measurements of static back muscle force and endurance are reliable and valid with our designed setup in older women with and without hyperkyphosis.

Corresponding Author: Minoo Khalkhali Zavieh, PT, PhD, m.khalkhali@sbmu.ac.ir.

Declaration of Interest
The authors report no declarations of interest.

Keywords

Spine; back extensor; maximal static force; endurance; hyperkyphosis; women; reliability; validity

INTRODUCTION

Hyperkyphosis is a common condition that often progresses with aging. Greater kyphosis is associated with impaired physical performance, and has a negative impact on health and quality of life (Chow and Harrison, 1987; Kado, 2009; Katzman et al, 2012; Katzman et al, 2013; Takahashi et al, 2005). Back extensor muscle strength and endurance are important in the preservation of normal spinal alignment and prevention of future spinal deformity (Mika, Unnithan, and Mika, 2005; O'Sullivan et al, 2006). Back extensor muscle strengthening interventions have reportedly improved age-related hyperkyphosis (Bansal, Katzman, and Giangregorio, 2014; Bautmans, Van Arken, Van Mackelenberg, and Mets, 2010; Greig, Bennell, Briggs, and Hodges, 2008). Thus, reliable and simple devices are needed for monitoring the effects of spinal muscle strength and endurance interventions on kyphosis in older adults.

Various analytic methods for testing back muscle force and endurance are described in the literature. Static measurements use pushing and pulling tests are the most common. The maximum static spinal extensor force, a measure of muscle strength (Chapman, 1970) is often assessed in a pushing test using hand-held dynamometer (Valentin and Maribo, 2014) and permanent strain-gauge installations (Mika, Unnithan, and Mika, 2005; Paalanne et al, 2009; Sinaki et al, 2005) in different testing positions and in diverse participant groups. Specialized dynamometers are also used in pushing tests for the measurement of trunk muscle torque (Granito et al, 2012; Smidt et al, 1983), however, all of these methods have their limitations. The assessment of static back muscle force using a hand-held dynamometer, in the prone position, has poor reliability (Moreland et al, 1997), because the examiner may not be able to provide a sufficient counter pressure against the participant's effort, especially in the prone position. Valentin and Maribo solved this problem with a hand-held dynamometer fixed with a tripod and a belt system. They reported strong reliability (ICC: 0.90) using this setup in prone lying for the assessment of maximal back extensor force (Valentin and Maribo, 2014). Participants of their study were older women with osteoporotic vertebral compression fractures, but the degree of thoracic kyphosis was not specifically reported. In other studies of older women, back muscle force has been assessed using a pushing test with a strain-gauge device (Mika, Unnithan, and Mika, 2005; Miyakoshi et al, 2005; Sinaki et al, 2005), but none of them investigated the reliability of their method. Moreover, these setups are permanent installations that have been primarily used for research and not integrated into clinical practice because of their high cost (Valentin and Maribo, 2014). Specialized equipment such as the MedX™ dynamometer (Graves et al, 1990) and isokinetic dynamometer (Karatas, Gogus, and Meray, 2002) showed strong reliability in normal participants, however they are expensive and time-consuming to perform.

Another method for testing static back muscle force and endurance is a pulling test. Previous studies described this type of testing to be practical (Jorgensen and Nicolaisen, 1986; Moreau, Green, Johnson, and Moreau, 2001) and not easily influenced by anthropometric factors (Jorgensen and Nicolaisen, 1986). In pulling tests of the back extensors, the participants sit (Bonde-Petersen, Mork, and Nielsen, 1975) or stand (Nicolaisen and Jorgensen, 1985; Singh, Bailey, and Lee, 2013) facing a strain-gauge dynamometer and attempt backward extension with the pelvis supported. The reliability of the pulling tests for back muscle force measurement has been moderate-high in healthy young populations (Jorgensen, 1997; Lariviere et al, 2001), however, the reliability of this type of test has not been tested in older women with or without hyperkyphosis.

Similar to maximal back muscle force, back muscle endurance has been assessed using several tests including: timed loaded standing (Shipp et al, 2000); the Sorensen test (Moffroid et al, 1994); isokinetic dynamometer test (Mayer, Gatchel, Betancur, and Bovasso, 1995); and electromyographic analysis (Kramer et al, 2005). Shipp et al. (2000) investigated the reliability of a combined measure of trunk and arm endurance using the timed loaded standing test in older women with and without vertebral fracture. This standing test measures the time a participant maintains both arms at 90 degrees shoulder flexion while holding 2-pound weights. They demonstrated good reliability in both groups (ICC > 0.8) (Shipp et al, 2000). They measured the degree of thoracic kyphosis in their participants, however, they assessed reliability of the endurance test in older women in general, not specifically those with hyperkyphosis. The Sorensen test, a test measuring the time the participant is able to extend their torso from a prone position on a table without support, may not be well tolerated because the participant does not receive feedback about the amount of force exerted (Moreau, Green, Johnson, and Moreau, 2001) and the prone test position is challenging for older adults with limited spinal mobility. Previous studies reported: back and leg pain (Biering-Sorensen, 1984); breathlessness (Moreland et al, 1997); and cramps of calves and discomfort in the head (Latikka, Battie, Videman, and Gibbons, 1995) which suggest intolerance to this test. Although electromyographic analysis of back extensor muscles is a safe technique, and a good indicator of muscle endurance, it is time-consuming and requires a trained examiner that is not generally available in a clinic. Pulling tests are another practical type of back muscle endurance measurement (Moreau, Green, Johnson, and Moreau, 2001). However, the reliability of the pulling tests has been reported in healthy, young populations (Jorgensen and Nicolaisen, 1986), but has not been investigated in older women.

Apart from reliability, the other important factor is the validity of device. The assessment of agreement between measurements by a target device and standard values is the most practical approach for evaluating instrument validity (Portney and Watkins, 2009). To ensure that a device is applicable in clinical practice, symptomatic participants should be included in validation studies of a new device (Weaver, Price, Czerniecki, and Sangeorzan, 2001), and the ability of the device to discriminate between healthy and symptomatic participants should be investigated (Pienaar and Barnard, 2017). None of previous studies reported instrument or construct validity of the device in older women with and without hyperkyphosis.

Due to the limited evidence about the reliability and validity of devices for measuring back extensor muscles force and endurance in older women with hyperkyphosis, we designed a pulling load cell for testing back extension force and endurance in a seated position with a simple and inexpensive device. Our first aim was to investigate the intra-rater reliability of a pulling load cell for measurement of maximum isometric back extensor force and endurance in older women with and without hyperkyphosis. Our secondary aim was to evaluate instrument validity and construct validity using known group methods for the outcomes of muscle force and endurance in older women with and without hyperkyphosis.

METHODS

Participants and Experimental Design

This study was carried out in a University biomechanics laboratory. Participants consisted of healthy, community-based older women (N=55), recruited through advertisements at urban entertainment districts. Inclusion criteria were 60– 80 years old, body mass index (BMI) of 25– 33 (the regional norm for BMI among women 60–64 years old) (Abbaszadeh Ahranjani et al, 2012), and the ability to stand and walk without an assistive device. Exclusion criteria included: cases of scoliosis; kyphoscoliosis; a history of back pain within the last year requiring medical attention; high blood pressure (more than 150/90); angina pectoris; spinal malignancy; or use of medications over the prior 12 months that could affect muscle performance (information was obtained from medical records). The study was approved by the Ethical Committee of Shahid Beheshti University of Medical Sciences. Prior to participation, each participant signed a written informed consent.

Eligible women were assigned at enrollment to either a hyperkyphosis (n= 30) group or normal (n = 25) group, based upon measurement of kyphosis using the Spinal Mouse (Mannion et al, 2004; Post and Leferink, 2004). Hyperkyphosis was defined as having a thoracic kyphosis of 50 degrees or higher (Granito et al, 2012; Sinaki et al, 2005). Six women with hyperkyphosis were excluded due inability to perform the test procedures correctly. Eleven women did not participate in day 2 testing session due to: low back pain after day 1 testing (4 in hyperkyphosis group, 7 in normal group) and prior to the initial testing session; and 5 participants withdrew from the study for personal reasons (1 in hyperkyphosis group, 4 in normal group). Finally, 33 healthy older women (hyperkyphosis group: n= 19, normal group: n= 14) completed the study and were included in the analysis (Figure 1).

Body Composition

Height (cm) using a standard stadiometer, and weight (kg) using a standard scale were collected at the initial visit. BMI was calculated (kg/m^2).

Static Back Extensor Force and Endurance Protocol

An “S” shape load cell (H3-C3–100 kg-3B-D55, Zemic, China) was used for measurement of maximum isometric force and endurance of back extensor muscles in a seated position (Figure 2). The setup included a stool with it’s anterior legs fixed to a wooden board, a fixed vertical bar and an “S” shape load cell. A lumbar back support with abdominal and pelvic restraints was used to stabilize the pelvis. The height of the load cell on the vertical bar could be adjusted according to the height of participants. This vertical bar was firmly fused to the wall at 3 points and two screws on the vertical bar allowed adjustment of the load cell to the height of the participants (Figure 2). A vest and four inflexible ropes connected the participant to the load cell. Two rows of rings sewn at thoracic and lumbar levels allowed the rope length and load cell alignment to be adjusted.

Test procedures were standardized and the setup was calibrated before each initial test and retest. Each test started 5 minutes (min) after a walking/back stretching warm-up. Participants were instructed to sit on the stool with the hips and knees flexed to 90 degree (Mika, Unnithan, and Mika, 2005), with the thighs parallel to the seat. Arms were crossed on the abdomen and the feet rested on the wooden board. Abdominal and pelvic restraints were placed over the abdomen and secured to the back support of the stool. Another strap (thigh restraint) was placed over the top of the thighs (Graves et al, 1990) at the level of ASIS and tightened to prevent any vertical or forward movement of the thighs or pelvis during the test. After positioning the participant on the stool in neutral upright sitting, the load cell was moved vertically on the bar and aligned with the superior border of the manubrium in midline. Four inflexible ropes were connected to the load cell hook and fixed to rings of the vest. The ropes were shortened individually according to participant-vertical bar distance and participant height (Figure 3).

The participants were instructed to gradually increase backward force over 1–2 s then exert maximal force for 3–5 s before gradually relaxing over 1–2 s (Figure 3). The participants received verbal encouragement and visual feedback (participants observed a digital display and were encouraged to increase the number displayed on the monitor) for achieving maximum force. Each warm up trial was followed by three successive maximum effort trials separated by 60 s rest periods (Mika, Unnithan, and Mika, 2005). The force generated was processed by an electronic measuring device and converted to a digital display. The maximum force was documented (kg). An additional trial was made if the third trial was more than 5% higher than either of the two preceding trials to ensure that the highest possible value had been achieved (Limburg et al, 1991). 50% of the maximum force was determined. After a 20– 30 min rest interval, participants performed a sustained contraction at 50% of the maximum force as their target. Participants received verbal and visual feedback on the digital display to maintain the sustained contraction at the target force. When the force could no longer be maintained above 90% of the target level, the test was stopped (Moreau, Green, Johnson, and Moreau, 2001). After verbal instruction and one warm up trial, three endurance trials, separated by 30 seconds rest, were performed and maximum time was recorded.

Reliability Procedure

Participants were tested at two visits within a 72-hour interval to allow time to recover from any residual fatigue or soreness that might have been associated with the tests. Before the initial test day, women participated in one session to familiarize themselves with the equipment and procedures (no practice involved). All measurements were performed by the same examiner experienced in using the setup of back force and endurance measurements.

On both test days, equipment and assessment procedures were the same. All tests were repeated on the 2nd testing day, at approximately the same time of day 1 for each participant. During the second test, the examiner was blinded to the test results of the previous day.

Validity Procedure

Due to inaccessibility to the gold standard Isokinetic dynamometer for instrument validity determination, we assessed the correlation between an external load (incremental calibrated weights (kg) and the load cell force (kg) (Pienaar and Barnard, 2017) using a rope-pulley system and standard weights. First, the weights were calibrated on a force plate. The load cell was fixed horizontally on a vertical bar and a rope-pulley system was embedded on another bar facing the load cell. One end of an inflexible rope was connected to the hook of the load cell and a standard weight was hung on the other end of rope, creating a pull force on the load cell, simulating the procedure of a back extensor force test. This procedure was repeated in 0.5 kg weight increments between 0.5 and 10 kg, and the load cell force was recorded from digital display.

Also, we assessed construct validity of the setup using known groups method (Portney and Watkins, 2009). On both groups, the outcomes of muscle force and endurance of test and retest were extracted. The mean force and endurance of each session were compared between older women with and without hyperkyphosis posture.

Statistical Analysis

Normality of distribution for all variables was determined using the Shapiro Wilk test in both groups ($P > 0.05$). Baseline demographic characteristics were presented as mean ± SD. To determine intra-tester reliability of the designed setup, a paired t-test was used to test systematic differences between day 1 testing and day 2 testing. Intra-class correlation (ICC) (two-way mixed and average 2 measurement was calculated (3.2) with a corresponding 95% confidence intervals. The criteria of Munro was used for interpretation of ICC: 0.00– 0.29 as very low correlation, 0.30– 0.49 as low correlation, 0.50– 0.69 as moderate correlation, 0.70– 0.89 as high correlation and 0.90– 1.00 as very high correlation (Munro and Visintainer, 2005). Standard error of measurement (SEM= SD of 1st test × (square root of (1-ICC)) and minimal detectable change (MDC= SEM × 1.96 × square root of 2)) for the back muscle tests were calculated.

Instrument validity was assessed using the Pearson correlation coefficient (r) to examine the correlation between load cell force (kg) and an applied external force (calibrated weights in kg) (Pienaar and Barnard, 2017). Bland-Altman plot with 95% limits of agreement was calculated to quantify the level of agreement between load cell force and external force (Giavarina, 2015). We used known groups method to support construct validity in a between-group comparison (Portney and Watkins, 2009). Mean muscle force and endurance of session 1 and 2 were compared using an independent t-test. For all statistical tests, a significance level of 0.05 was used. All analyses were performed with SPSS software package (version 16.0).

RESULTS

The study sample consisted of two groups of healthy, community-dwelling older women, 19 with hyperkyphosis (mean kyphosis 55.5 (SD: 6.4) degrees) and 14 with a normal curvature (mean kyphosis 39.5 (SD: 5.3) degrees). The mean age was 67.0 (SD: 5.0) years in the hyperkyphosis group and 63.0 (SD: 6.0) years in the normal group (Table 1).

The mean for all measurements on both testing days and the results of paired t-tests are presented in Table 2. For back extensor muscle force, the ICC of the hyperkyphosis and normal group was 0.96 and 0.97 respectively (Table 3). The SEM for maximal back extensor force was 1.4 kg for the hypkyphosis group and 1.0 kg for the normal group. The MDC was 3.9 kg for the hypkyphosis group and 2.7 kg for the normal group. The ICC of back extensor muscle endurance was 0.82 in the hyperkyphosis and 0.89 in the normal group. The SEM for back extensor endurance was 27.0 s in the hyperkyphosis group and 24.9 seconds in the normal group. The MDC for back extensor endurance was 74.9 s in the hyperkyphosis group and 69.2 seconds in the normal group (Table 3).

There was a high correlation between load cell force and applied external force ($r = 0.99$, $P < 0.001$). Figure 4 illustrates the results of a Bland-Altman plot with the average of the forces (load cell force and applied external weight) plotted against the absolute difference between the two forces. On both days, the hyperkyphosis group showed significantly lower mean back extensor force ($P = 0.001$ (day 1), $P < 0.001$ (day 2)) and endurance ($P = 0.001$ (day 1), $P < 0.001$ (day 2)) than the normal group (Table 4).

DISCUSSION

The present study evaluated the reliability and validity of the pulling load-cell setup for measuring static back muscle force and endurance, and demonstrated that the designed load-cell setup is reliable when repeated by the same tester assessing maximal back extensor force and endurance in women with hyperkyphosis and normal kyphosis. Instrument validity was demonstrated by a high correlation between load cell and calibrated external forces. We also found differences in the muscle force and endurance between the two groups which supports the known group validity, and ability of the designed setup to discriminate between groups for the outcomes of maximal back extensor force and endurance. Pulling tests are simple, easy to use and low cost (Moreau, Green, Johnson, and Moreau, 2001), but the reliability and validity of them has not been previously tested in older populations. Our study was the first study to examine the reliability and construct validity of a mounted pulling test to assess static spinal extensor muscle force and endurance in older women with and without hyperkyphosis.

Our results confirmed the reliability results of pulling tests in younger adults (Demoulin et al, 2012; Essendrop, Schibye, and Hansen, 2001). The reliability of back extensor force testing was greater than 0.95 and the values of SEM and MDC were low in both groups ($SEM < 1.5$, $MDC < 4$ kg) in our study (Table 3). Our results are also better than a previous study by Valentin and Maribo (2014) comparing intra-rater reliability of two measurement methods of static back extensor force in older osteoporotic women in the prone position.

They reported better reliability using a hand-held dynamometer fixed with a tripod (ICC: 0.90, SEM: 20.5 (N), smallest detectable change SDC, considered the same as MDC: 56.7 (N)) compared with a hand-held dynamometer fixed by the tester (ICC: 0.75, SEM: 27.1 (N), SDC: 75.2 (N)). This difference may have several explanations. Valentin and Maribo (2014) enrolled older women (mean age: 72 (9.3)) with osteoporotic vertebral compression fractures. Our participations were healthy older women (mean age of 67 (5)). The average back muscle force at session 1 in the study by Valentin and Maribo (2014) was approximately 136 N compared to approximately 28 kg (274 N) in our hyperkyphosis group. Therefore, it is possible that the different results in the two studies may be participant, not setup related. On the other hand, there are methodological differences in the test position and the amount of external fixation during the tests. In Valentin and Maribo (2014) study, the participant was positioned in prone lying on a couch, with the hips and knees in a neutral position without external fixation. It is also possible that it was difficult for participants to maintain sufficient contact between their back and the dynamometer in the prone position testing. We tested participants in a seated position which is usually better tolerated by many persons with spinal deformities (Mika, Unnithan, and Mika, 2005). We also used a belt for stabilizing the thigh and pelvis, and we provided a familiarization session before the initial test which may have reduced measurement error (Essendrop, Schibye, and Hansen, 2001). Another difference between the two studies is the calculation of SEM. Valentine and Maribo used the formula “SD/square root of 2” (Hopkins, 2000), and we used the formula “SEM = SD* square root of 1-ICC” (Weir, 2005) which limits the ability to compare the SEM results of two studies. To overcome this problem, we recalculated SEM using their formula, and the values of SEM in both of our groups were lower than the Valentin and Maribo (2014) study (hyperkyphosis group: 1.86 kg, normal group: 1.06 kg). Even though the measurement unit was different in both studies, SEM is dimensionless and these values of reliability of measurement may be compared (Hopkins, 2000). According to these calculations, it is possible that our setup had better reliability, although given our small sample size, we cannot accurately make this definitive conclusion. Furthermore, while the hand-held dynamometer has been considered a reliable and valid tool for assessing muscle force in extremities (Stark et al, 2011), the utility of this device for testing back extensor force is questionable (Moreland et al, 1997). Besides the challenges of prone positioning with the hand-held dynamometer, it is hard to be consistent between examiners (Moreland et al, 1997). Another study investigated inter-rater reliability of the hand-held dynamometer for measurement of static back extensor force in 39 healthy younger workers. This study reported poor inter-rater reliability (ICC: 0.24, SEM: 68 N), indicating the challenges in consistency of this setup between raters (Moreland et al, 1997).

The intra-tester ICC for back extensor endurance was 0.82 in hyperkyphosis group (SEM: 27.0 seconds, MDC:74.9 seconds) and 0.89 in normal group (SEM: 24.9 seconds: MDC: 69.2 seconds), but it is hard to compare this to other measurements of back extensor endurance due to so much variation in back extensor endurance testing methodology and the populations tested. In previously described endurance tests such as the prone double leg raise (McIntosh, Wilson, Affieck, and Hall, 1998) and Sorensen (Moffroid et al, 1994) tests, there is an unknown amount force exerted by the participant. For healthy participants and participants with LBP, test-retest reliability of the Sorensen and modified Sorensen tests ranged from 0.54 to 0.99 (Moreau, Green, Johnson, and Moreau, 2001). One study of young participants, using a hand held dynamometer in position similar to Sorensen test, showed moderate reliability of back extensor endurance (ICC= 0.59, SEM= 20 s) (Moreland et al, 1997). However, the Sorensen has not been investigated in older women with hyperkyphosis, and it could be detrimental and have an adverse effect on performance.

Our values of test-retest ICC were similar to a previous study (Shipp et al, 2000) who developed the Timed Loaded Standing test, a combined measure of trunk and arm endurance, performed standing holding 2-pound weights in each hand with the arms at 90 degrees of shoulder flexion. The back extensor endurance test in a standing position is considered to be more appropriate than the Sorensen test among older adults because it is less sensitive to heterogeneous physiques and performed in a more comfortable standing position (da Silva et al, 2005; Kankaanpaa et al, 1998). Same day, intra-rater, within session (inter-trial) ICC was 0.89 in the group without vertebral fractures and 0.81 in the group with vertebral fractures. Test-retest (6–10 day) ICC (mean of both trials) was 0.84 in the group without vertebral fractures and 0.85 in the group with vertebral fractures (Shipp et al, 2000). Although both tests are simple (i.e. the TLS and our designed load cell) our method focused on the measurement of back extensor endurance and proved to be a reliable setup. It appears that the designed setup in our study is another good option for the measurement of back extensor endurance in older women.

The Biodex dynamometer is another alternative for testing back extensor endurance in a sitting position (Moreau, Green, Johnson, and Moreau, 2001), however the Biodex system is both space-consuming and high cost (Moreau, Green, Johnson, and Moreau, 2001). Van Dieen and Heijblom determined the test retest reliability of Biodex system in healthy young participants performing a sustained contraction at 50% of the maximum until no longer able to maintain above a 90% target level. Test retest ICC between sessions was 0.54 and within session (after a 5 min rest of the first test) was 0.94 (van Dieen and Heijblom, 1996). We designed a low cost and low tech setup (approximately 260 USD) with minimal space requirements, and our setup showed better reliability than the Biodex in the back endurance test. Nevertheless the small sample size in our study limits the ability to make definitive recommendations for clinical use. Future study with a large sample of older women with hyperkyphosis is needed.

The importance of SEM and MDC is to help clinicians separate true change from measurement error (Johnson et al, 2012), although there is disagreement about the best method among researchers. Atkinson and Nevill (1998) argue that SEM underestimates true change and suggests MDC should be used instead because MDC accurately provides an index of difference between measurements (Weir, 2005). According to the SEM results of the current study, in older women with hyperkyphosis following an exercise program, an increase in maximal back extensor force of more than 1.4 kg and endurance time of more than 27.0 s should be considered real change. If we use values of MDC, an increase more than 3.9 kg and 74.9 s are a therapeutic effect. In comparison with SEM, MDC is more conservative, and may result in clinicians ignoring clinical changes (Hopkins, 2000).

According to the results of the Pearson correlation and Bland-Altman plot, we believe that instrument validity was supported. While we did not have access to a gold standard isokinetic dynamometer, we assessed the instrument validity by examining the correlation between the documented load cell force (kg) and the applied calibrated weights (kg), and report a strong and statistically significant correlation between the two forces (r = 0.99). Furthermore, the degree of agreement between the two forces determined by a Bland-Altman plot was also strong. The middle line illustrating the mean of the difference between the two forces is near zero.

We assessed the construct validity of the setup using the known group method and found evidence of validity that the designed setup and measurement procedure discriminates between the two groups for the outcomes of maximal back extensor force and endurance. Except one study (Shipp et al, 2000), none of the previous studies investigated validity of their setup, so we are unable to make a comparison. Shipp et al. (2000) assessed the concurrent validity of TLS by examining the association between TLS and measures of physical impairments, functional performance and functional status among women with vertebral fracture. Except for thoracic kyphosis and weight, the results of their validity study showed statistically significant correlations between TLS and measurements. Also, women with vertebral fracture were categorized into 2 groups according to back pain (yes or no), and independent t-test determined there were differences in mean TLS time between the two groups (Shipp et al, 2000). Although their results demonstrated acceptable concurrent validity, the mean TLS time was not compared between women with and without vertebral fracture. Thus, the ability of the TLS to discriminate between healthy (no fracture) and symptomatic (fracture) participants is not known. We found statistically significant differences in the muscle force and endurance between the two groups which suggests that the designed setup has sufficient ability to detect low back extensor force and endurance in women with hyperkyphosis.

The decrease in back extensor force and endurance is a common problem associated with aging. In order to assess these changes, simple, reliable and valid measurements are needed. Our designed load cell is a reliable, valid and cost-effective device for measuring back extensor force and endurance. The values of SEM and MDC may help clinicians and researchers monitor true changes in the performance of back extensor muscles. The clinical use of this setup for assessment of static back extensor force and endurance in older women with and without hyperkyphosis should investigated in larger samples in future studies.

Limitations

Our study has several limitations. One important limitation is high dropout rate. 6 out of 29 women with hyperkyphosis we’re not able to complete the test procedures correctly but did not have pain and 11 of 44 older women were unable to complete day 2 testing procedures due to pain. We excluded participants with a history of low back pain, but incorrect performance of the tests may have resulted in low back pain. We suggest using a practice trial in the familiarization session and improving the instructions to ensure slow buildup of force. These modifications may lead to decreased dropout rate. Another limitation is the small sample size that limits our ability to make definitive recommendations for clinical use of this setup for back muscle force and endurance measurements. These results need to be confirmed in a large sample with different thoracic kyphosis degrees. We did not investigate inter-tester reliability, because participants would have to repeat the testing by two testers on two days, and we considered this an excessive burden to our participants. However, we did contribute data on intra-rater reliability which is a first step towards demonstrating utility of this setup. Due to lack of evidence about the reliability of pulling tests for measuring of back extensor muscle in older adults, we were not able to compare ICCs of the current study with that of other published results. Also, different study population and different statistical methods from previous studies in older women prevent definitive conclusions about the preferred setup. Therefore, our results cannot be generalized to the overall population of older adults, although we have now contributed data that can be useful for future age-grouped comparisons. Because we did not have access to the gold standard isokinetic dynamometer for comparison, we were unable to evaluate criterion-related validity. Future work is needed to address this. Even though age-related hyperkyphosis and associated back extensor muscle weakness affects both sexes, our study included women only. Future studies should be undertaken to include the reliability and validity of back extensor force and endurance tests in older men.

CONCLUSION

Our study reports high intra-rater reliability and validity of the designed load cell setup for measurement of back extensor muscle force and endurance in older women with and without hypkyphosis. We determined the standard error of measurement and minimal detectable change for all measurements, and this information can be used by clinicians and researchers when assessing back extensor muscles after an intervention. Future studies are needed to assess the reliability and criterion-related validity of our designed load cell with large samples of participants of both sexes, and among individuals with different degrees of thoracic kyphosis.

ACKNOWLEDGEMENTS

This study is related to the project NO 1395/66075 from Student Research Committee, Shahid Beheshti University of Medical Sciences, Tehran, Iran. We also appreciate the “Student Research Committee” and “Research & Technology Chancellor” in Shahid Beheshti University of Medical Sciences for their financial support of this study. Authors thanks to Motahare Ababaf Behbahani (MSC student) for assisting with measurements and data collections. This work was supported by the National Institute on Aging (NIA) (grant number: R01 AG041921); and the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (grant number: NIH P50 AR063043).

REFERENCES

Abbaszadeh Ahranjani SH, Kashani H, Forouzanfar MH, Meybodi HA, Larijani B, Aalaa M, Mohajeri-Tehrani MR 2012 Waist circumference, weight, and body mass index of Iranians based on National Non-Communicable Disease Risk Factors Surveillance. Iranian Journal of Public Health 41 35–45. [PubMed: 23113163]

Atkinson G, Nevill AM 1998 Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine 26: 217–238. [PubMed: 9820922]

Bansal S, Katzman WB, Giangregorio LM 2014 Exercise for improving age-related hyperkyphotic posture: A systematic review. Archives of Physical Medicine and Rehabilitation 95: 129–140. [PubMed: 23850611]

Bautmans I, Van Arken J, Van Mackelenberg M, Mets T 2010 Rehabilitation using manual mobilization for thoracic kyphosis in elderly postmenopausal patients with osteoporosis. Journal of Rehabilitation Medicine 42: 129–135. [PubMed: 20140408]

Biering-Sorensen F 1984 Physical measurements as risk indicators for low-back trouble over a one-year period. Spine 9: 106–119. [PubMed: 6233709]

Bonde-Petersen F, Mork AL, Nielsen E 1975 Local muscle blood flow and sustained contractions of human arm and back muscles. European Journal of Applied Physiology and Occupational Physiology 34: 43–50. [PubMed: 1149731]

Chapman AE 1970 The measurement of muscular strength. British Journal of Sports Medicine 5: 42–49.

Chow RK, Harrison JE 1987 Relationship of kyphosis to physical fitness and bone mass on post-menopausal women. American Journal of Physical Medicine 66: 219–227. [PubMed: 3434624]

da Silva RA, Arsenault AB, Gravel D, Lariviere C, de Oliveira E 2005 Back muscle strength and fatigue in healthy and chronic low back pain subjects: A comparative study of 3 assessment protocols. Archives of Physical Medicine and Rehabilitation 86: 722–729. [PubMed: 15827924]

Demoulin C, Grosdent S, Smeets R, Verbunt J, Jidovtseff B, Mahieu G, Crielaard JM, Vanderthommen M 2012 Muscular performance assessment of trunk extensors: A critical appraisal of the literature. In: Asghar Norasteh A (Ed) Low Back Pain, pp 141–166. INTECH Open Access Publisher.

Essendrop M, Schibye B, Hansen K 2001 Reliability of isometric muscle strength tests for the trunk, hands and shoulders. International Journal of Industrial Ergonomics 28: 379–387.

Giavarina D 2015 Understanding Bland Altman analysis. Biochemia Medica 25: 141–151. [PubMed: 26110027]

Granito RN, Aveiro MC, Renno AC, Oishi J, Driusso P 2012 Comparison of thoracic kyphosis degree, trunk muscle strength and joint position sense among healthy and osteoporotic elderly women: a cross-sectional preliminary study. Archives of Gerontology and Geriatrics 54: e199–202. [PubMed: 21831460]

Graves JE, Pollock ML, Carpenter DM, Leggett SH, Jones A, MacMillan M, Fulton M 1990 Quantitative assessment of full range-of-motion isometric lumbar extension strength. Spine 15: 289–294. [PubMed: 2141187]

Greig AM, Bennell KL, Briggs AM, Hodges PW 2008 Postural taping decreases thoracic kyphosis but does not influence trunk muscle electromyographic activity or balance in women with osteoporosis. Manual Therapy 13: 249–257. [PubMed: 17433756]

Hopkins WG 2000 Measures of reliability in sports medicine and science. Sports Medicine 30: 1–15. [PubMed: 10907753]

Johnson KD, Kim KM, Yu BK, Saliba SA, Grindstaff TL 2012 Reliability of thoracic spine rotation range-of-motion measurements in healthy adults. Journal of Athletic Training 47: 52–60. [PubMed: 22488230]

Jorgensen K 1997 Human trunk extensor muscles physiology and ergonomics. Acta Physiologica Scandinavica Supplementum 637: 1–58. [PubMed: 9246395]

Jorgensen K, Nicolaisen T 1986 Two methods for determining trunk extensor endurance. A comparative study. European Journal of Applied Physiology and Occupational Physiology 55: 639–644. [PubMed: 2946578]

Kado DM 2009 The rehabilitation of hyperkyphotic posture in the elderly. European Journal of Physical and Rehabilitation Medicine 45: 583–593. [PubMed: 20032918]

Kankaanpaa M, Laaksonen D, Taimela S, Kokko SM, Airaksinen O, Hanninen O 1998 Age, sex, and body mass index as determinants of back and hip extensor fatigue in the isometric Sorensen back endurance test. Archives of Physical Medicine and Rehabilitation 79: 1069–1075. [PubMed: 9749686]

Karatas GK, Gogus F, Meray J 2002 Reliability of isokinetic trunk muscle strength measurement. American Journal of Physical Medicine and Rehabilitation 81: 79–85. [PubMed: 11807340]

Katzman W, Cawthon P, Hicks GE, Vittinghoff E, Shepherd J, Cauley JA, Harris T, Simonsick EM, Strotmeyer E, Womack C, Kado DM 2012 Association of spinal muscle composition and prevalence of hyperkyphosis in healthy community-dwelling older men and women. Journal of Gerontology. Series A, Biological Sciences and Medical Sciences 67: 191–195.

Katzman WB, Huang MH, Lane NE, Ensrud KE, Kado DM 2013 Kyphosis and decline in physical function over 15 years in older community-dwelling women: The Study of Osteoporotic Fractures. Journal of Gerontology. Series A, Biological Sciences and Medical Sciences 68: 976–983.

Kramer M, Ebert V, Kinzl L, Dehner C, Elbel M, Hartwig E 2005 Surface electromyography of the paravertebral muscles in patients with chronic low back pain. Archives of Physical Medicine and Rehabilitation 86: 31–36. [PubMed: 15640986]

Lariviere C, Gagnon D, Gravel D, Bertrand Arsenault A, Dumas J, Goyette M, Loisel P 2001 A triaxial dynamometer to monitor lateral bending and axial rotation moments during static trunk extension efforts. Clinical Biomechanics 16: 80–83. [PubMed: 11114447]

Latikka P, Battie MC, Videman T, Gibbons LE 1995 Correlations of isokinetic and psychophysical back lift and static back extensor endurance tests in men. Clinical Biomechanics 10: 325–330. [PubMed: 11415575]

Limburg PJ, Sinaki M, Rogers JW, Caskey PE, Pierskalla BK 1991 A useful technique for measurement of back strength in osteoporotic and elderly patients. Mayo Clinic Proceedings 66: 39–44. [PubMed: 1824867]

Mannion AF, Knecht K, Balaban G, Dvorak J, Grob D 2004 A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: Reliability of measurements and comparison with data reviewed from the literature. European Spine Journal 13: 122–136. [PubMed: 14661104]

Mayer T, Gatchel R, Betancur J, Bovasso E 1995 Trunk muscle endurance measurement. Isometric contrasted to isokinetic testing in normal subjects. Spine 20: 920–927. [PubMed: 7644957]

McIntosh G,Wilson L, Affieck M, Hall H 1998 Trunk and lower extremity muscle endurance: normative data for adults. Journal of Rehabilitation Outcomes Measurement 2: 20–39.

Mika A, Unnithan VB, Mika P 2005 Differences in thoracic kyphosis and in back muscle strength in women with bone loss due to osteoporosis. Spine 30: 241–246. [PubMed: 15644764]

Miyakoshi N, Hongo M, Maekawa S, Ishikawa Y, Shimada Y, Okada K, Itoi E 2005 Factors related to spinal mobility in patients with postmenopausal osteoporosis. Osteoporosis International 16: 1871–1874. [PubMed: 15959615]

Moffroid M, Reid S, Henry SM, Haugh LD, Ricamato A 1994 Some endurance measures in persons with chronic low back pain. Journal of Orthopaedic and Sports Physical Therapy 20: 81–87. [PubMed: 7920605]

Moreau CE, Green BN, Johnson CD, Moreau SR 2001 Isometric back extension endurance tests: A review of the literature. Journal of Manipulative and Physiological Therapeutics 24: 110–122. [PubMed: 11208223]

Moreland J, Finch E, Stratford P, Balsor B, Gill C 1997 Interrater reliability of six tests of trunk muscle function and endurance. Journal of Orthopaedic and Sports Physical Therapy 26: 200–208. [PubMed: 9310911]

Munro BH, Visintainer MA 2005 Statistical Methods for Health Care Research Philadelphia, PA, Lippincott Williams and Wilkins.

Nicolaisen T, Jorgensen K 1985 Trunk strength, back muscle endurance and low-back trouble. Scandinavian Journal of Rehabilitation Medicine 17: 121–127. [PubMed: 2932794]

O'Sullivan PB, Mitchell T, Bulich P, Waller R, Holte J 2006 The relationship beween posture and back muscle endurance in industrial workers with flexion-related low back pain. Manual Therapy 11: 264–271. [PubMed: 15953751]

Paalanne NP, Korpelainen R, Taimela SP, Remes J, Salakka M, Karppinen JI 2009 Reproducibility and reference values of inclinometric balance and isometric trunk muscle strength measurements in Finnish young adults. Journal of Strength and Conditioning Research 23: 1618–1626. [PubMed: 19620899]

Pienaar AW, Barnard JG 2017 Development, validity and reliability of a new pressure air biofeedback device (PAB) for measuring isometric extension strength of the lumbar spine. Journal of Medical Engineering and Technology 41: 216–222. [PubMed: 27838945]

Portney LG, Watkins MP 2009 Foundations of Clinical Research: Application to Practice (3rd ed). Upper Saddle River, New Jersey: Pearson/Prentice-Hall.

Post RB, Leferink VJ 2004 Spinal mobility: Sagittal range of motion measured with the Spinal Mouse, a new non-invasive device. Archives of Orthopaedic and Trauma Surgery 124: 187–192. [PubMed: 14968367]

Shipp KM, Purse JL, Gold DT, Pieper CF, Sloane R, Schenkman M, Lyles KW 2000 Timed loaded standing: a measure of combined trunk and arm endurance suitable for people with vertebral osteoporosis. Osteoporosis International 11: 914–922. [PubMed: 11193243]

Sinaki M, Brey RH, Hughes CA, Larson DR, Kaufman KR 2005 Balance disorder and increased risk of falls in osteoporosis and kyphosis: Significance of kyphotic posture and muscle strength. Osteoporosis International 16: 1004–1010. [PubMed: 15549266]

Singh DK, Bailey M, Lee R 2013 Decline in lumbar extensor muscle strength the older adults: Correlation with age, gender and spine morphology. BMC Musculoskeletal Disorders 14: 215. [PubMed: 23875830]

Smidt G, Herring T, Amundsen L, Rogers M, Russell A, Lehmann T 1983 Assessment of abdominal and back extensor function. A quantitative approach and results for chronic low-back patients. Spine 8: 211–219. [PubMed: 6222489]

Stark T, Walker B, Phillips JK, Fejer R, Beck R 2011 Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. PM&R 3: 472–479. [PubMed: 21570036]

Takahashi T, Ishida K, Hirose D, Nagano Y, Okumiya K, Nishinaga M, Matsubayashi K, Doi Y, Tani T, Yamamoto H 2005 Trunk deformity is associated with a reduction in outdoor activities of daily living and life satisfaction in community-dwelling older people. Osteoporosis International 16: 273–279. [PubMed: 15235766]

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 International 25: 2143–2149. [PubMed: 24866393]

van Dieen JH, 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. [PubMed: 8618156]

Weaver K, Price R, Czerniecki J, Sangeorzan B 2001 Design and validation of an instrument package designed to increase the reliability of ankle range of motion measurements. Journal of Rehabilitation Research and Development 38: 471–475. [PubMed: 11732825]

Weir JP 2005 Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research 19: 231–240. [PubMed: 15705040]

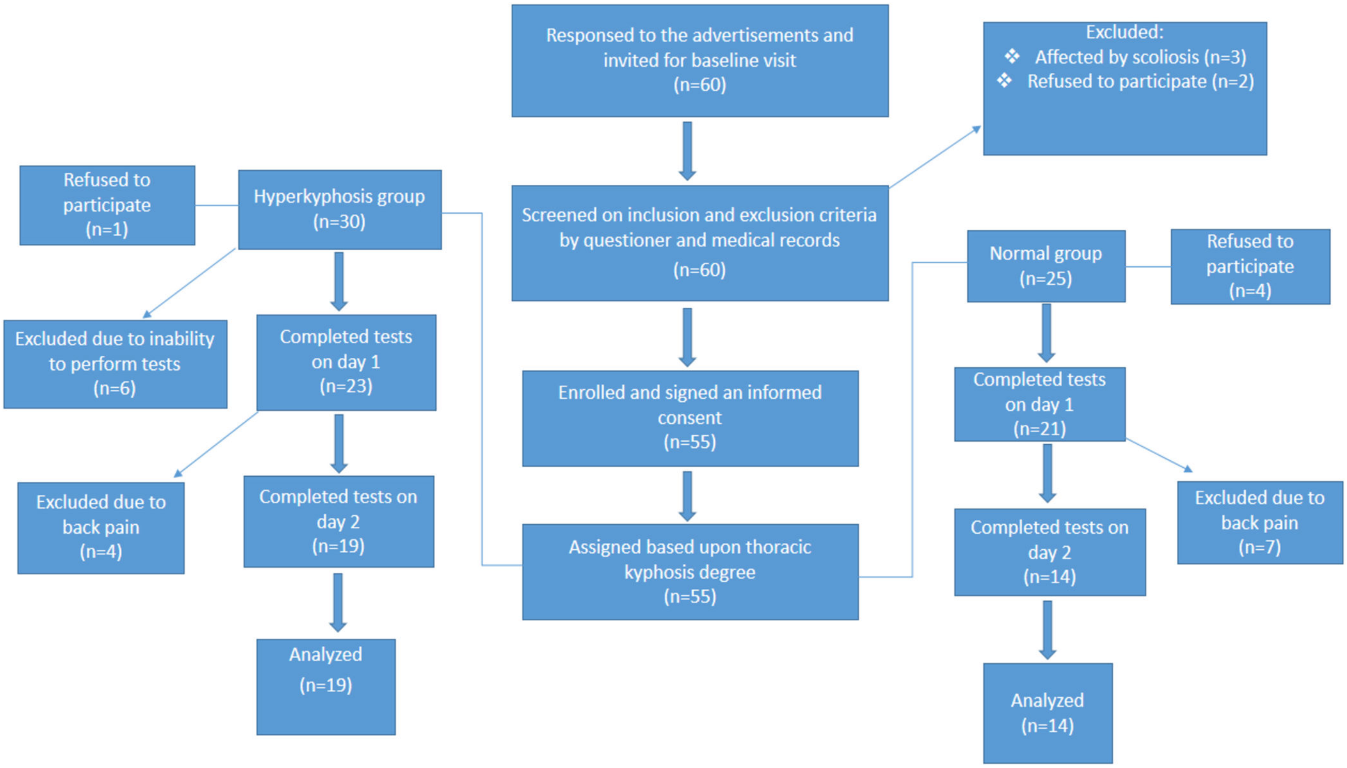

Figure 1.

Flow chart of study participation enrollment, screening, categorized and analysis.

Figure 2.

The designed setup for back extensor muscle tests. a. Vertical bar, b. An" S" shape load cell, c. Wooden board, d. Lumbar back support, e. Two screw for adjustment of load cell, f. Hook for connection of vest & rope to the load cell, g. monitor.

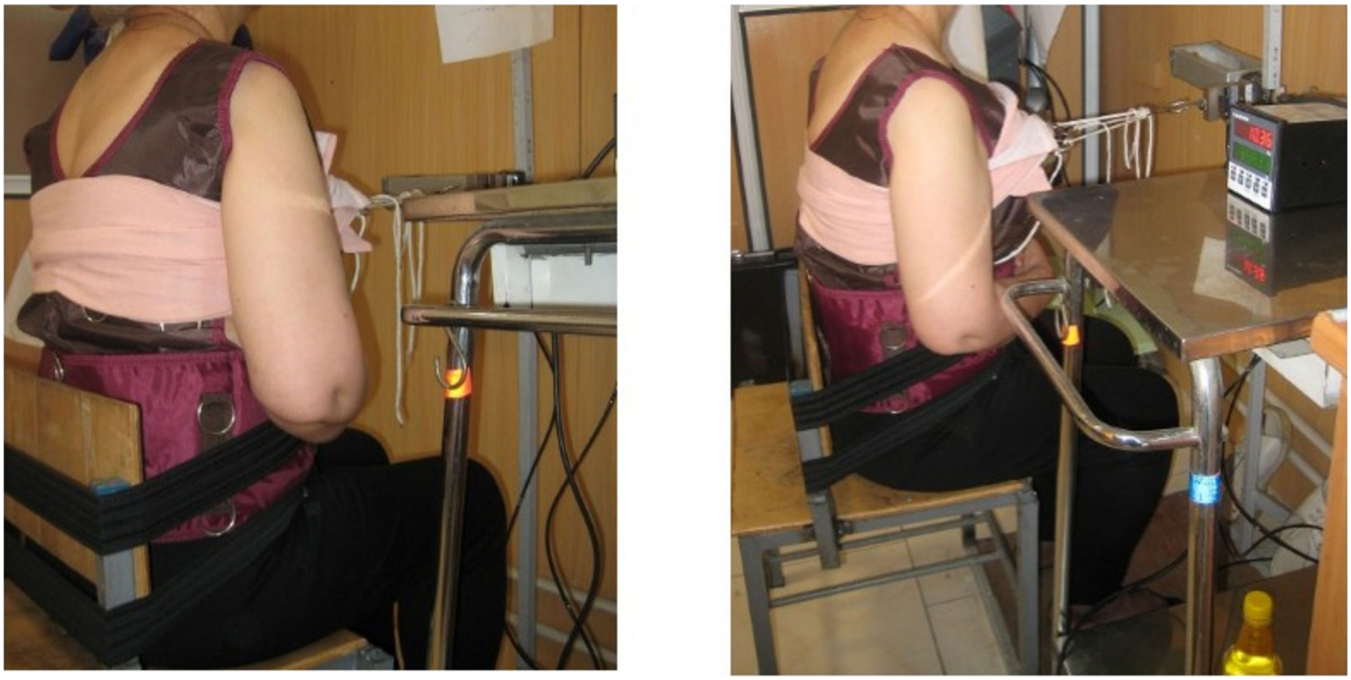

Figure 3.

Assessing the maximum static back extensor force using the designed setup in sitting position. Subject sit on the stool facing testing apparatus, abdominal, pelvic restraints were placed over the abdomen and secured to the back support of the stool and thigh restraint was placed over the top of the thighs. The load cell was aligned with the superior border of the manubrium in midline. Four inflexible ropes were connected to the load cell hook and fixed to rings of the vest. Subject pull the trunk back as hard as they could under control of examiner.

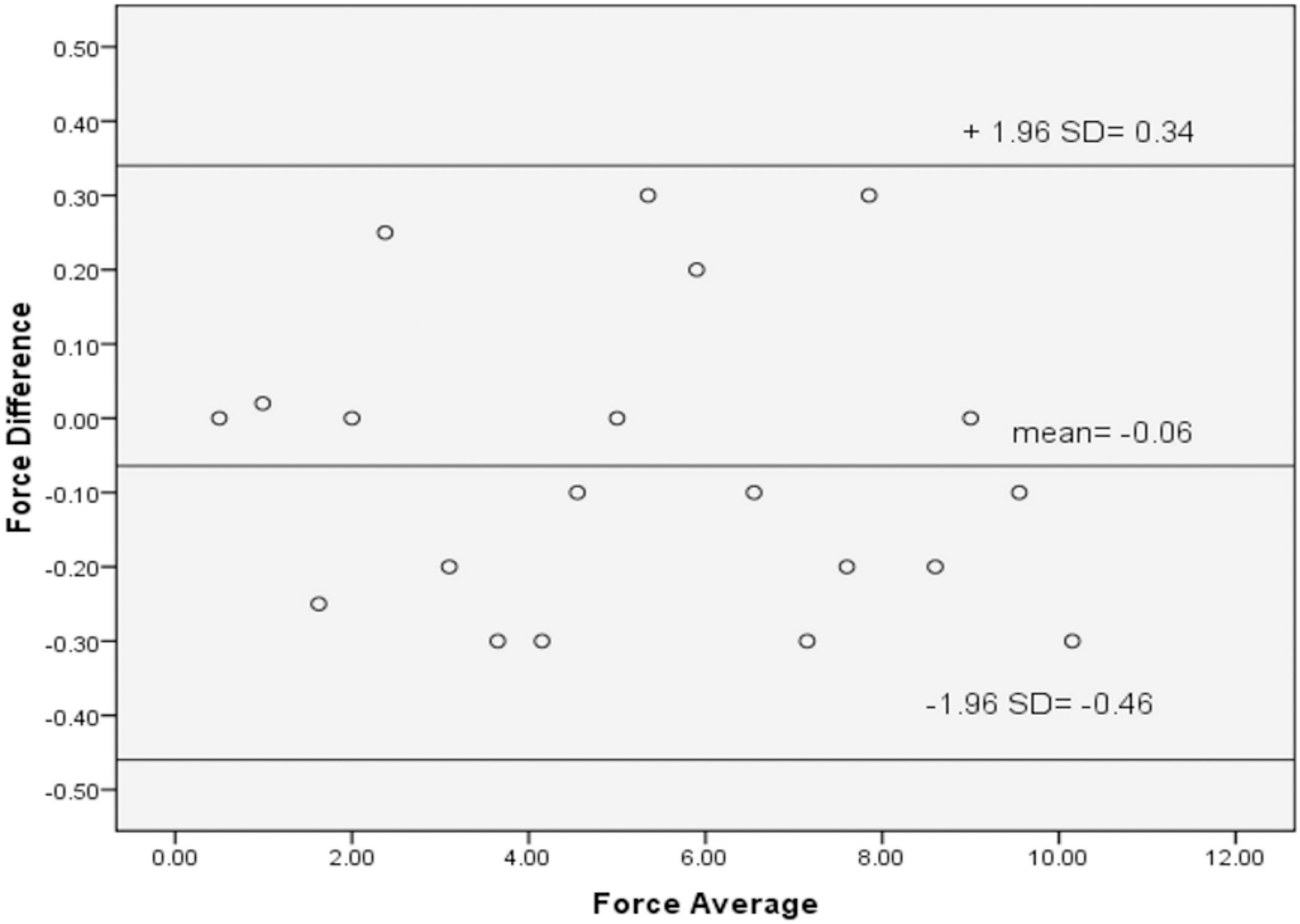

Figure 4.

Bland–Altman plot provided no evidence of heteroscedasticy since the scatter of differences is uniform across the average of two forces.

Table 1.

Baseline demographic characteristics of the 33 subjects (mean ± SD).

Variable	**Hyperkyphosis group (N= 19)**	**Normal group (N= 14)**
Age (yr)	67.0 ± 5.0	63.0 ± 6.0
Height (cm)	156.0 ± 4.8	153.0 ± 5.5
Weight (kg)	69.3 ± 7.7	62.7 ± 9.8
BMI (kg/m^2)	28.5 ± 3.4	26.7 ± 3.4
Thoracic kyphosis (^{0})	55.5 ± 6.4	39.5 ± 5.3
Lumbar lordosis (^{0})	−35.3 ± 9.8	−33.0 ± 8.1

Abbreviations: yr: years; BMI: body mass index; cm: centimeter; kg: kilogram; kg/m^2: kilogram/meter^2; ^{0}: degree.

Table 2.

Descriptive statistics for back extensor force and endurance on 2 different days.

Variable	**Hyperkyphosis group (N=19)**			**Normal group (N= 14)**		
	Day 1	**Day 2**	**p value** [a]	**Day 1**	**Day 2**	**p value** [a]
	Mean (SD)	**Mean (SD)**		**Mean (SD)**	**Mean (SD)**	
Back extensor force (kg)	28.2 (7.0)	27.5 (7.0)	0.23	36.6 (5.7)	37.7 (5.6)	0.07
Back extensor endurance (s)	153.0 (63.8)	160.3 (52.0)	0.48	238.0 (75.3)	244.6 (56.0)	0.56

Abbreviations: SD: standard deviation; kg: kilogram; s: second.

[a] P value of paired-T-test comparing day1 and day 2

Table 3.

Intrarater reliability of back extensor muscle tests.

Variable	**Hyperkyphosis group (N=19)**			**Normal group (N=14)**		
	ICC (95% CI)	**SEM**	**MDC**	**ICC (95%CI)**	**SEM**	**MDC**
Back extensor force (kg)	0.96 (0.90– 0.98)	1.4	3.9	0.97 (0.92–0.99)	1.0	2.7
Back extensor endurance (s)	0.82 (0.56–0.93)	27.0	74.9	0.89 (0.68–0.96)	24.9	69.2

Abbreviations: ICC: Intraclass correlation coefficient; CI: confidence interval; SEM: Standard error of measurement; MDC: minimal detectable change; kg: kilogram; s: second.

Table 4.

Comparison of variables, hyperkyphosis versus normal group.

Variable	Day 1			Day 2		
	Hyperkyphosis group (N=19)	Normal group (N=14)	p value [a]	Hyperkyphosis group (N=19)	Normal group (N=14)	p value [a]
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
Back extensor force (kg)	28.2 (7.0)	36.6 (5.7)	0.001	27.5 (7.0)	37.7 (5.6)	P < 0.001
Back extensor endurance (s)	153.0 (63.8)	238.0 (75.3)	0.001	160.3 (52.0)	244.6 (56.0)	P < 0.001

Abbreviations: SD: standard deviation; kg: kilogram; s: second.

[a] significant difference between groups.