Original article

Isokinetic trunk flexion–extension protocol to assess trunk muscle strength and endurance: Reliability, learning effect, and sex differences

María Pilar García-Vaquero, David Barbado, Casto Juan-Recio, Alejandro Lopez-Valenciano, Francisco J. Vera-Garcia *

Sports Research Centre, Miguel Hernandez University of Elche, Elche, Alicante, Spain

Purpose: The purpose of this study was to examine the reliability and the learning effect of an isokinetic trunk flexion–extension protocol designed to simultaneously assess trunk muscle strength and endurance. In addition, the effect of the participants' sex on the reliability data was examined.

Methods: Fifty-seven healthy and physically active young men ($n = 28$) and women ($n = 29$) performed the isokinetic protocol 5 times, separated by a week between each of the first 4 sessions and by a month between the last 2 sessions. The protocol consisted of performing 4 trials of 15 maximum flexion–extension concentric exertions at 120°/s (range of trunk motion = 50°). The absolute and relative peak torque and total work were calculated to assess trunk flexion and extension strength. In addition, endurance ratio, modified endurance ratio, fatigue final ratio, recovery ratio, and modified recovery ratio variables were used for the assessment of trunk muscle endurance in both directions.

Results: Regarding the absolute reliability, no relevant changes were found between paired-comparison sessions for most strength and endurance variables, except for total work and relative total work variables in the flexion movement in both sexes. In addition, the typical error of the isokinetic variables was lower than 10% in both males and females, and minimum detectable changes ranged from 7%–20%, with a tendency to be higher in females and in endurance variables. The strength variables showed high intraclass correlation coefficients (ICC; >0.74); however, for the endurance variables only the endurance ratio and the modified endurance ratio obtained moderate to high ICC values ($0.57 < ICC < 0.82$). In addition, the analysis of the variance reported no significant differences between consecutive pairs of sessions for most variables in both sexes.

Conclusion: Overall, these findings provide clinicians, trainers, and researchers with a 10-min single-session protocol to perform a reliable muscle strength and endurance evaluation of trunk flexor and extensor muscles, all within the same protocol.

© 2016 Production and hosting by Elsevier B.V. on behalf of Shanghai University of Sport. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Core muscles; Fitness; Isokinetic dynamometry; Performance; Spine; Testing

1. Introduction

The contribution of the trunk musculature to many sports (e.g., taekwondo, judo, tennis, golf, baseball, handball, rowing, *etc.*)[1–5] and daily life activities has aroused considerable interest in trainers, clinicians, and researchers.[6,7] In the field of sports, it is thought that increases in the ability to exert the maximum trunk muscle force (trunk muscle strength), as well as the ability to exert trunk muscle force repeatedly or continuously over a long period of time (trunk muscle endurance), can improve athletic performance[1,8,9] and help prevent and treat back disorders in individuals with trunk muscle weakness.[4,10]

For these reasons, many field and laboratory protocols have been developed to assess trunk muscle strength and endurance in sports, fitness, clinical, and research settings.

For decades, isokinetic dynamometry has been widely used to measure trunk muscle strength in sports performance[1,11,12] as well as to identify injury risks[13,14] and to assess the progress of rehabilitation programs[15,16] in clinical settings. The main reasons for its popularity are the validity and reliability shown by the isokinetic instrument,[17] the relative and absolute reliability of the isokinetic strength protocols,[18–21] and the ability to measure different muscle groups while controlling contraction type, angular velocity, range of motion, body position, number of repetitions and sets, *etc.*[22] In addition, because previous studies have not found a learning effect for these protocols,[18–21,23–26] participants do not have to carry out a long period of practice before testing.

In contrast, trunk muscle endurance has normally been evaluated using field tests[27–30] because they are easy to perform, do not require large and expensive equipment, and allow numerous people to be evaluated all at once in a short period of time. However, several researchers have questioned their use, especially in the field of sports, for several reasons: (1) the lack of specificity of some protocols to trunk demands for a particular sport;[7,8,31] (2) the influence of individual anthropometry[32] and test practice/experience[28] on scores; and (3) the large absolute reliability of most field protocols,[32–34] which brings into question their ability to detect real improvement in the athletic population.[35] Based on the isokinetic dynamometry characteristics presented here (i.e., instrumental reliability, performance control, nonlearning effect, *etc.*), isokinetic trunk endurance protocols could be an alternative to field tests; however, to the best of our knowledge there are few studies on isokinetic trunk endurance,[1,36,37] and only the study by Mayer et al.[37] has assessed protocol reliability. In this study, two different trunk muscle strength and endurance protocols were analyzed, and only the strength variables showed high reliability, whereas the reliability of the endurance variables was considerably lower.

Taking into account the lack of isokinetic trunk endurance protocols and the time constraints in sports and clinical settings, which make performing several protocols difficult, an isokinetic trunk flexion–extension protocol was developed to simultaneously evaluate both trunk muscle strength and trunk muscle endurance. The protocol was based on those developed by Mayer et al.[37] and has a short execution time (approximately 10 min), which facilitates its use in professional and scientific fields. Although this protocol has recently been used to show the contribution of trunk muscle function to high-level performance in judo,[1] its reliability has not been analyzed. Therefore, the main purpose of this study was to assess the absolute and relative reliability and the learning effect of this new isokinetic trunk flexion–extension protocol. In addition, we examined the effect of the participants' sex on the reliability data, because there are only a few studies on isokinetic trunk dynamometry that do not show consistent results that have evaluated protocol reliability depending on the participants' sex.[23–25]

2. Methods

2.1. Participants

Fifty-seven healthy young volunteers, 28 males (age: 24.1 ± 3.3 years; height: 176 ± 5.2 cm; mass: 75.4 ± 8.6 kg) and 29 females (age: 22.2 ± 3.8 years; height: 164.1 ± 4.8 cm; mass: 59.0 ± 7.1 kg), took part in this study. They were physically active, performing 1–3 h of moderate physical activity 1–3 days per week. Participants, who were recruited from the university population, took part in a variety of recreational physical activities such as team sports, aerobic exercises, and strength workout routines, but none of them was involved in trunk strength and/or endurance training programs. They completed a questionnaire about their medical and sports history to assess their health status and regular physical activity. None of the participants reported a recent history of back injury, abdominal surgery, or inguinal hernia, and all participants were free of neurologic, cardiorespiratory, or musculoskeletal disorders. All subjects were informed of the risks of this study and signed an informed consent based on the 2013 Declaration of Helsinki, which was approved by the Ethics Committee of the university.

2.2. Testing protocol description

The isokinetic trunk protocol was performed on a Biodex isokinetic dynamometer (Model 2000, System 4 Pro; Biodex Corporation, Shirley, NY, USA). Participants were placed on the dual-position back extension–flexion attachment of the dynamometer with the trunk upright, the hips and knees flexed at 90°, the thighs parallel to the floor, and the dynamometer axis of rotation aligned with the imaginary line joining the anterior superior iliac spines.[38] This was considered the anatomic reference position (Fig. 1). To hold the participant to the dynamometer attachment, adjustable pads were placed behind the head, the sacrum, and the upper trunk and on the anterior surface of the tibia; in addition, Velcro straps were placed on the upper trunk, the thighs, and the pelvis. The trunk range of movement was limited at 50°, with 30° (−30°) of trunk flexion and 20° (+20°) of trunk extension, relative to the anatomic reference
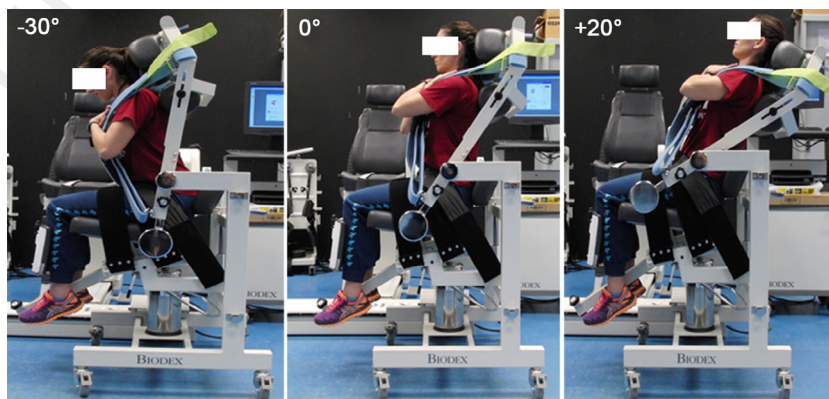

Fig. 1. Participant performing a maximum effort of trunk flexion–extension in the isokinetic dynamometer with a range of motion of 50° (−30° trunk flexion; 0° initial position; 20° trunk extension).

position (0°) described earlier (Fig. 1). According to Grabiner et al.,[38] ranges of trunk motion no larger than 50° isolate lumbar motion, reducing hip flexion–extension. Moreover, the location of the dynamometer axis of rotation at the anterior superior iliac spine level and the use of the pad behind the sacrum and the strap on the pelvis minimized hip motion during the protocol.

The protocol consisted of 4 sets of 15 consecutive maximum concentric trunk flexion and extension efforts with 1 min rest between sets.[1] It started from the flexion position and was performed with an angular velocity of 120°/s. This angular velocity was chosen because it is considered to be safe and reliable for measuring mechanical work.[39] Participants were told to keep their hands and arms crossed over their chest during the protocol. In addition, they were instructed to perform the maximum effort from the beginning of the first set and to maintain it until the end of the test. Moreover, they were verbally encouraged with the same indications and intensity across repetitions to exert maximum physical effort throughout the protocol.[20,40]

Before testing, participants carried out a warm-up that consisted of 1 set of 10 submaximum trunk flexion–extension exertions at testing angular velocity (120°/s). This warm-up period helped participants become familiar with the equipment and test execution. The overall testing duration was approximately 10 min.

Taking into account that at least 3 administrations of a protocol are needed to estimate its reliability accurately,[41] each participant executed 5 testing sessions of the isokinetic trunk flexion–extension protocol. All the trials were performed at the same time of the day and were managed by the same researcher. For each participant, the position on the dynamometer was recorded in a log sheet during the first testing session and was controlled across sets (adjusting pads and straps) and testing sessions to ensure protocol reliability.[19–21,40] There was a 1-week rest period between the first, second, third, and fourth testing session. However, because a weekly 1-min trunk training session has shown to be effective to improve trunk flexor endurance in adolescents with no experience in trunk exercise programs,[42] a 1-month rest was given between the fourth and fifth testing sessions to examine the possible influence of a training effect on the reliability analysis.

2.3. Data reduction

Fig. 2 shows an example of the force time-history for the isokinetic trunk protocol. The first 3 repetitions of each set were discarded to avoid nonreal maximum executions related to the beginning of the isokinetic performance, because most participants reached their maximum strength values after the fourth repetition (82.9% and 72.9% for the extension and flexion movement, respectively). Therefore, 12 repetitions per set (i.e., the fourth to fifteenth) were considered for further analysis.

The absolute (raw scores) and relative (scores divided by body mass) peak torque (N·m) and relative peak torque (N·m/kg) and the absolute and relative total work (TW) obtained from the entire set (in joules (J)) and relative total

work (in J/kg) were calculated for each set. Considering that most participants did not achieve the maximum strength values during the first set, especially for the extension movement (75% of the participants), the strength values obtained in the 2 best sets were averaged for each variable and direction to assess trunk flexion and extension strength.

In addition, 5 variables were used for the assessment of trunk muscle endurance in both directions (expressed in percentages):

1. Endurance ratio (ER), obtained after dividing the work (W) performed during the last 3 repetitions of each set by the W performed during the fourth, fifth, and sixth repetition of each set and multiplied by 100.[37]

$$ER = \frac{\sum W(13,14,15)}{\sum W(4,5,6)} \times 100$$

2. Modified endurance ratio (MER), obtained after dividing the W performed during the last 3 repetitions of each set by 3 times the maximum W (MW) reached in any repetition during the set and multiplied by 100.

$$MER = \frac{\sum W(13,14,15)}{3 \times MW(\text{rep.})} \times 100$$

3. Fatigue final ratio (FFR), obtained after dividing the W performed during the last 3 repetitions in the last set by 3 times the maximum W performed in any repetition of any set and multiplied by 100.[1,37]

$$FFR = \frac{\sum W(13,14,15)(\text{set } 4)}{3 \times MW(\text{rep.})(\text{sets})} \times 100$$

4. Recovery ratio (RR), obtained after dividing the TW performed during the last set by the TW performed during the first set and multiplied by 100.[37]

$$RR = \frac{TW(\text{set } 4)}{TW(\text{set } 1)} \times 100$$

5. Modified recovery ratio (MRR), obtained after dividing the TW performed during the last set by the maximum TW (maxTW) performed in any set and multiplied by 100.

$$MRR = \frac{TW(\text{set } 4)}{\text{maxTW}(\text{sets})} \times 100$$

Notice that ER and MER represent the ability to maintain the force output throughout each set, whereas final fatigue ratio, recovery ratio, and modified recovery ratio represent the ability to maintain the force output between sets. Therefore, a lower value for these variables represents a higher drop in trunk muscle force throughout the repetitions and/or sets, that is, a lower endurance score. Because many participants did not show a force decrement during the first set (mainly in extension
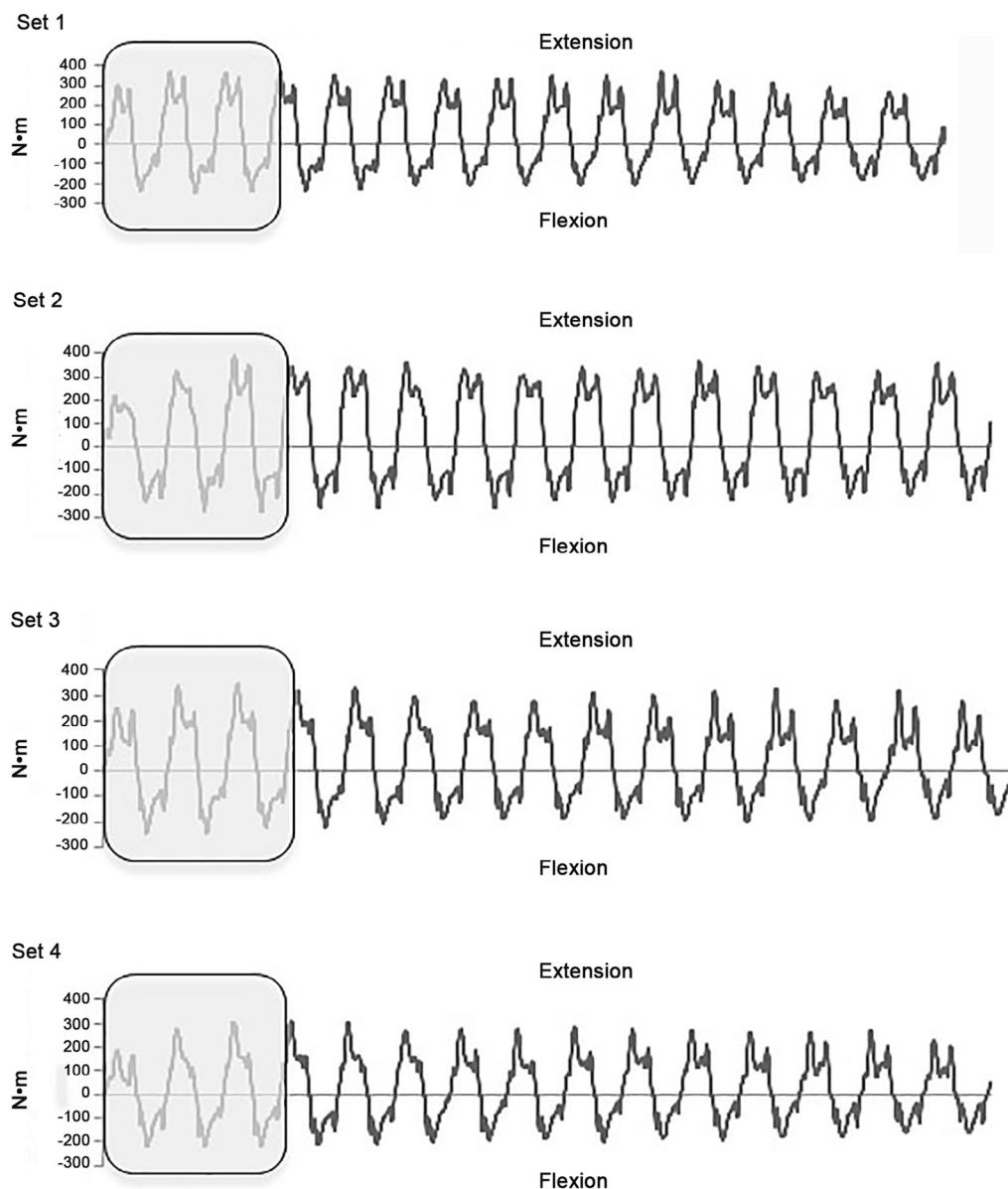
4   M.P. García-Vaquero et al.

Fig. 2. Force time-history of a participant for the isokinetic protocol (4 sets × 15 repetitions). As has been explained in section 2.3, the first three repetitions (*shaded*) were not used for the data analysis.

direction) for ER and MER, the 3 sets with the lowest scores were averaged for further analyses.

2.4. Statistical analyses

The distribution of raw data sets was checked using the Kolmogorov–Smirnov test, which demonstrated that all data had a normal distribution ($p > 0.05$). Descriptive statistics including means and standard deviations were calculated separately for each variable for both males and females. Briefly, a 9 (4 strength variables and 5 endurance variables) × 5 (testing sessions) × 2 (males and females) analysis of variance with repeated measures in the last factor was used to identify score differences between sessions (i.e., learning effect). When significant differences were obtained, *post hoc t* test analyses with Bonferroni adjustments were performed for multiple comparisons. Mauchly's test was used to check the assumption of sphericity of the data.

The detection of a possible heteroskedasticity of the random error distribution between paired sessions was done with calculation of Pearson's correlation coefficient (r) between absolute individual test–retest differences and individual means of each consecutive pair of sessions. No significant correlations showed the absence of heteroskedasticity, so raw data were used for the statistical analyses.[43,44]

To analyze the intersession absolute reliability of each variable, the typical error (TE; within-subject variation) and the change in the mean (between consecutive pairs of sessions) with their respective 90% confidence limits and the minimum detectable change (MDC; $1.5 \times$ TE) were calculated using the method previously described by Hopkins[39] and Hopkins

et al.[41,45] The absolute reliability was calculated to average the reliability for the consecutive pairs of trials (2–1, 3–2, 4–3, 5–4). The TE was established using the following formula: SDdiff/√2, where SDdiff is the standard deviation of the difference between consecutive pairs of sessions. The change in the mean was calculated as the mean difference between consecutive pairs of sessions. For the change in the mean, the probability that the true value of the effect was positive, trivial, or negative was inferred as follows: most unlikely, <0.5%; very unlikely, 1%–5%; unlikely, 5%–25%; possibly, 25%–75%; likely, 75%–95%; very likely, 95%–99%; and most likely, >99%.[45] The current study considered a "relevant or substantial" change when a change between paired-comparison sessions was statistically significant ($p > 0.05$) and the probability of the worthwhile differences was "likely" or higher (>75%; positive or negative).

The relative reliability of the different measures was analyzed using the intraclass correlation coefficient ($ICC_{2,1}$), calculating 90% confidence limits. According to Hopkins[39] and Hopkins et al.,[41,45] the ICC was calculated from the analysis of variance $(F - 1)/(F + k - 1)$, in which F is the F ratio for the subject term and k is the number of trials (i.e., 2). The ICC values were categorized as follows: excellent (0.90–1.00), high (0.70–0.89), moderate (0.50–0.69), and low (<0.50).[46]

Statistical analysis was performed with the SPSS statistics software (version 18.0 for Windows 7; SPSS Inc., Chicago, IL, USA), establishing significance as $p \leq 0.05$.

3. Results

Descriptive statistics and the change in the mean between consecutive testing sessions for the isokinetic strength and endurance variables are displayed in Tables 1 and 2, respectively. For the strength variables in males and females, the change in the mean was generally above "likely trivial," except for specific cases shown in Table 1. Furthermore, the analysis of variance with repeated measures indicated no significant interaction effect among sessions for the extension movement in any variable or either sex. In contrast, in the flexion movement a few slightly significant differences were found (Table 1). These differences were mainly detected when we compared session 1 with the rest of the sessions, because it showed higher strength values (Fig. 3).

For the endurance variables (Table 2), the changes in the mean were mainly "possibly trivial" for both males and females in flexion and extension movements, except for a few cases shown in Table 2. In addition, the analysis of variance with repeated measures reported no significant differences between consecutive pairs of sessions.

Test–retest reliability statistics for the isokinetic strength and endurance variables (between consecutive pairs of testing sessions (2–1, 3–2, 4–3, 5–4)) are presented in Tables 3 and 4, respectively. To facilitate result comprehension, data have been presented as the mean of the 4 paired testing sessions.

The strength variables (Table 3) showed mean ICC values above 0.74 (0.74 < ICC < 0.91) and mean TE values below 10% (5% < TE < 10%), whereas minimal detectable change was lower than 14% (8% < MDC < 14%). Comparing males and
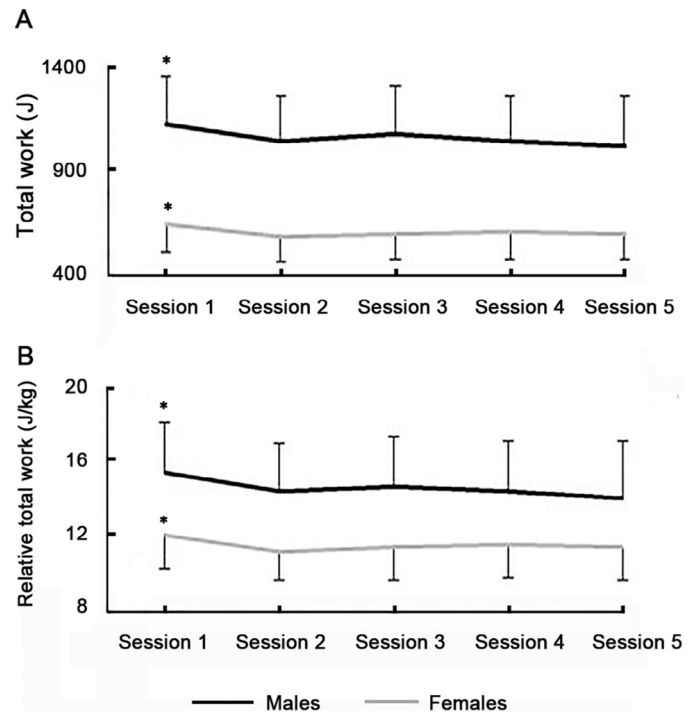

Fig. 3. Evolution of absolute (A) and relative (B) total work throughout the 5 sessions of the study for flexion direction. Asterisks over or under the error bars mean significant differences between sessions 1 and 2. J, joule.

females, similar relative and absolute reliability values were obtained.

The endurance variables (Table 4) showed lower relative reliability values compared with strength variables, especially in females and in trunk flexion movement. However, endurance ratio and modified endurance ratio for trunk flexion and extension movements and final fatigue ratio for trunk extension movement in males, and endurance ratio and modified endurance ratio for trunk extension movement in females, presented moderate to high mean ICC scores (0.57 < ICC < 0.82).

Regarding absolute reliability, most variables showed mean values of TE below 10% (4% < TE < 10%), but final fatigue ratio and recovery ratio in males and females and modified recovery ratio in females presented mean values that ranged from 11%–14% (Table 4). Minimal detectable change ranged from 7%–20%, tending to be higher in females than in males.

4. Discussion

Although isokinetic dynamometry protocols are commonly used to assess trunk muscle strength in clinical and sports fields, they have seldom been used to evaluate trunk muscle endurance.[1,36,37] The purpose of this study was to examine the reliability and the learning effect of an isokinetic protocol designed to simultaneously assess trunk muscle strength and endurance in physically active males and females. The main findings of the current study were the high and moderate relative reliability for the strength and endurance variables, respectively. Thus, both variables seem adequate for ranking individuals according to their strength or endurance level.[41,47] In

6   M.P. García-Vaquero et al.

Table 1

Descriptive values (mean ± standard deviation (SD)) for testing session 1, the change in the mean (ChM) between consecutive testing sessions (mean change ± 90% confidence limits (CL)) and their probabilistic inference about the true magnitude of change for the *isokinetic strength variables* in males and females.

Variable	Session 1 Mean ± SD	Session 2–Session 1 ChM ± CL	Inference	Session 3–Session 2 ChM ± CL	Inference	Session 4–Session 3 ChM ± CL	Inference	Session 5–Session 4 ChM ± CL	Inference
Males									
PT (N·m)									
Extension	373.02 ± 60.10	1.94 ± 12.20	Likely trivial	10.02 ± 9.77	Likely trivial	−11.78 ± 11.49	Possibly trivial	1.70 ± 9.16	Most likely trivial
Flexion	227.12 ± 25.25	−5.83 ± 5.08	Possibly negative	2.76 ± 5.73	Likely trivial	0.85 ± 6.69	Likely trivial	−4.23 ± 6.14	Possibly trivial
RPT (N·m/kg)									
Extension	4.98 ± 0.77	0.01 ± 0.16	Likely trivial	0.14 ± 0.13	Possibly trivial	−0.11 ± 0.13	Possibly trivial	0.03 ± 0.12	Very likely trivial
Flexion	3.05 ± 0.44	−0.09 ± 0.07	Possibly trivial	0.04 ± 0.08	Likely trivial	0.01 ± 0.09	Likely trivial	−0.05 ± 0.08	Likely trivial
TW (J)									
Extension	2139.32 ± 501.25	11.06 ± 112.44	Likely trivial	115.02 ± 77.83	Possibly positive	−159.81 ± 71.96	Likely negative	0.27 ± 81.11	Very likely trivial
Flexion	1147.34 ± 209.54	−78.04 ± 43.10	Likely negative*	28.14 ± 49.52	Possibly trivial	−19.00 ± 29.34	Likely trivial	−22.69 ± 33.17	Likely trivial
RTW (J/kg)									
Extension	28.48 ± 6.25	0.04 ± 1.50	Likely trivial	1.52 ± 1.06	Possibly positive	−2.16 ± 0.97	Likely negative	0.07 ± 1.12	Likely trivial
Flexion	15.32 ± 2.78	−1.05 ± 0.54	Likely negative*	0.37 ± 0.65	Possibly trivial	−0.37 ± 0.43	Possibly trivial	−0.29 ± 0.46	Likely trivial
Females									
PT (N·m)									
Extension	249.44 ± 41.97	6.56 ± 8.76	Possibly trivial	−3.08 ± 9.81	Likely trivial	−4.97 ± 7.03	Likely trivial	8.42 ± 8.32	Possibly trivial
Flexion	173.03 ± 22.30	−9.15 ± 5.99	Likely negative	2.31 ± 5.72	Likely trivial	−5.16 ± 6.13	Possibly negative	−2.17 ± 5.47	Likely trivial
RPT (N·m/kg)									
Extension	4.24 ± 0.64	0.10 ± 0.15	Possibly trivial	−0.05 ± 0.17	Likely trivial	−0.08 ± 0.12	Likely trivial	0.13 ± 0.14	Possibly trivial
Flexion	2.96 ± 0.42	−0.16 ± 0.11	Likely negative	0.05 ± 0.10	Likely trivial	−0.08 ± 0.11	Possibly trivial	−0.03 ± 0.09	Likely trivial
TW (J)									
Extension	1435.26 ± 323.58	69.73 ± 69.59	Possibly positive	−32.39 ± 55.31	Likely trivial	−46.82 ± 49.74	Likely trivial	31.87 ± 63.41	Possibly trivial
Flexion	701.09 ± 125.96	−50.91 ± 14.43	Most likely negative*	13.40 ± 20.70	Likely trivial	5.01 ± 19.55	Likely trivial	−8.66 ± 21.67	Likely trivial
RTW (J/kg)									
Extension	24.28 ± 4.68	1.20 ± 14.43	Possibly positive	−0.53 ± 20.70	Possibly trivial	−0.79 ± 19.55	Possibly trivial	0.38 ± 21.67	Possibly trivial
Flexion	11.90 ± 1.81	−0.87 ± 0.25	Most likely negative*	0.25 ± 0.37	Possibly trivial	0.09 ± 0.32	Likely trivial	−0.14 ± 0.34	Likely trivial

Notes: Terms for chances: possibly, 25%–75%; likely, 75%–95%; very likely, 95%–99%; most likely, >99%.

* Significance: $p \leq 0.05$.

Abbreviations: PT = peak torque; RPT = relative peak torque; RTW = relative total work; TW = total work.

addition, strength and endurance variables showed low absolute reliability values, indicating that they may be useful in detecting real changes when an intervention (treatment or training) is applied.[41,47] Finally, significant improvements in strength and endurance variables were not found across sessions, suggesting that a single testing session could be enough to assess trunk muscle strength and endurance.

Regarding the relative reliability, isokinetic strength variables in flexion and extension efforts showed high ICC values in both males and females (0.74 < ICC < 0.91) (Table 3). These findings agree with previous studies in which the strength was measured in different isokinetic conditions (velocity, range of motion, isokinetic devices, subject placement, *etc.*).[18,20,21,25,37,40,48] Overall, the results of all these studies indicate the robustness of isokinetic measures in assessment of trunk muscle strength.

Concerning endurance variables, we found moderate to high ICC values for those variables that assessed the drop in the performance within sets (endurance ratio and modified

Table 2
Descriptive values (mean ± standard deviation (SD)) for testing session 1, the change in the mean (ChM) between consecutive testing sessions (mean change ± 90% confidence limits (CL)), and their probabilistic inference about the true magnitude of change for the *isokinetic endurance variables* in males and females.

Variable	Session 1 Mean ± SD	Session 2–Session 1		Session 3–Session 2		Session 4–Session 3		Session 5–Session 4	
		ChM ± CL	Inference	ChM ± CL	Inference	ChM ± CL	Inference	ChM ± CL	Inference
Males									
ER (%)									
Extension	89.98 ± 12.51	−1.22 ± 2.22	Likely trivial	−0.08 ± 2.17	Likely trivial	0.89 ± 2.10	Possibly trivial	−0.30 ± 2.71	Likely trivial
Flexion	82.63 ± 6.33	0.29 ± 1.98	Possibly trivial	−0.23 ± 2.43	Possibly trivial	−2.34 ± 1.89	Possibly negative	−0.90 ± 1.92	Possibly trivial
MER (%)									
Extension	81.02 ± 7.41	−0.37 ± 1.84	Likely trivial	0.82 ± 1.74	Possibly trivial	0.34 ± 1.96	Likely trivial	0.34 ± 1.90	Likely trivial
Flexion	77.99 ± 5.95	−0.28 ± 1.63	Possibly trivial	−0.42 ± 1.93	Possibly trivial	−1.49 ± 1.54	Possibly negative	0.18 ± 1.95	Possibly trivial
FFR (%)									
Extension	71.30 ± 14.31	−0.34 ± 3.67	Likely trivial	−0.39 ± 2.97	Likely trivial	3.38 ± 3.25	Possibly positive	−1.22 ± 2.44	Likely trivial
Flexion	62.38 ± 10.10	3.83 ± 4.14	Likely positive	0.51 ± 3.36	Possibly trivial	−2.18 ± 2.58	Possibly negative	−1.19 ± 3.23	Possibly trivial
RR (%)									
Extension	95.96 ± 18.53	−6.06 ± 6.06	Possibly negative	−0.84 ± 4.57	Possibly trivial	4.65 ± 4.55	Likely positive	−2.52 ± 3.23	Possibly negative
Flexion	81.31 ± 13.16	6.28 ± 5.07	Likely positive	−1.42 ± 3.80	Possibly trivial	1.69 ± 4.52	Possibly positive	−3.87 ± 4.58	Possibly negative
MRR (%)									
Extension	87.79 ± 11.25	−0.70 ± 4.02	Possibly trivial	−2.11 ± 3.18	Possibly negative	2.94 ± 3.24	Possibly positive	−0.71 ± 2.28	Possibly trivial
Flexion	79.99 ± 10.72	4.31 ± 3.06	Likely positive	−1.18 ± 2.73	Possibly negative	1.35 ± 2.89	Possibly trivial	−3.28 ± 3.16	Likely positive
Females									
ER (%)									
Extension	71.30 ± 14.31	−3.39 ± 2.41	Likely negative	1.13 ± 2.18	Possibly trivial	0.02 ± 2.07	Likely trivial	−0.62 ± 1.87	Likely trivial
Flexion	62.38 ± 10.10	−2.45 ± 3.11	Possibly negative	0.31 ± 2.66	Possibly trivial	2.34 ± 2.61	Possibly positive	−2.26 ± 2.87	Possibly negative
MER (%)									
Extension	82.82 ± 6.42	−1.59 ± 1.46	Possibly negative	0.18 ± 2.14	Likely trivial	−0.26 ±1.87	Possibly trivial	−0.73 ± 1.99	Most likely trivial
Flexion	74.91 ± 6.95	−2.70 ± 3.31	Possibly negative	0.19 ± 2.60	Possibly trivial	1.04 ± 2.66	Possibly trivial	−2.17 ± 2.87	Possibly negative
FFR (%)									
Extension	76.67 ± 11.77	−2.59 ± 2.18	Possibly negative	0.84 ± 3.16	Possibly trivial	−1.10 ±2.82	Possibly trivial	−0.14 ± 2.83	Possibly trivial
Flexion	63.04 ± 11.63	−3.87 ± 3.82	Possibly negative	−0.15 ± 3.78	Likely trivial	2.16 ± 2.67	Possibly positive	−2.91 ± 4.02	Possibly negative
RR (%)									
Extension	101.70 ± 18.70	−7.15 ± 4.94	Likely negative	1.07 ± 4.59	Possibly trivial	−0.43 ±3.31	Possibly trivial	−1.32 ± 4.39	Possibly trivial
Flexion	80.14 ± 12.05	−0.12 ± 4.76	Possibly trivial	1.90 ± 3.66	Possibly trivial	0.59 ± 4.40	Possibly trivial	−0.10 ± 6.78	Possibly trivial
MRR (%)									
Extension	91.77 ± 8.62	−2.42 ± 1.91	Possibly negative	1.14 ± 3.08	Possibly trivial	0.15 ± 2.42	Possibly trivial	−0.76 ± 2.48	Possibly negative
Flexion	78.60 ± 9.95	−1.82 ± 3.47	Possibly trivial	2.39 ± 3.10	Possibly trivial	1.78 ± 3.67	Possibly trivial	−1.12 ± 5.33	Most likely trivial

Notes: Terms for chances: possibly, 25%–75%; likely, 75%–95%; very likely, 9%5–99%; most likely, >99%.
*Significance: $p \leq 0.0$.
Abbreviations: ER = endurance ratio; FFR = final fatigue ratio; MER = modified endurance ratio; MRR = modified recovery ratio; RR = recovery ratio.

endurance ratio (0.57 < ICC < 0.82)), mainly for flexion–extension movements in males and for extension movements in females (Table 4). However, those endurance variables that evaluated the drop in the performance between sets (final fatigue ratio, recovery ratio, and modified recovery ratio) obtained low relative reliability values (Table 4). It is possible that the rest time between sets was enough to allow some participants to partially recover from the effort performed, reducing or avoiding the drop in strength between sets in these participants. In the same way, some participants may have adopted conservative strategies during the protocol, not performing a maximum effort from the beginning of the protocol, which can be seen in some participants by the lack of a drop-off in work performance.[37,40] In general, the ICC values obtained in this study were higher than those found by Mayer et al. using similar variables (0.35 < ICC < 0.42),[37] maybe because in our protocol participants performed 4 sets and in Mayer et al.'s protocol participants performed 2 sets.

Regarding the absolute reliability, overall, strength, and endurance variables showed typical percentage errors close to or below 10% in both males and females (Tables 3 and 4).

Although there are no clear guidelines to decide the adequate cutoff that ensures the precision of the measurement, some authors have suggested that a variability of a measure lower than 10% could be considered appropriate for clinical and research purposes.[43,49] Therefore, most strength and endurance variables analyzed in this study seem to have good test–retest absolute consistency. To the best of our knowledge, no previous studies have examined the absolute reliability of isokinetic trunk endurance protocols. In relation to the isokinetic strength protocols, we found similar[18–21] or better[23,24,26] absolute reliability than previous studies did, which could be due to the fact that the angular velocities[23] and ranges of motion[18,21] used in some of the previous studies were higher than those used in our protocol (120°/s and 50° trunk motion). In this sense, angular velocities higher than 120°/s could increase the error between sessions,[23] and large ranges of motion could result in a misalignment between the biological axis of the trunk and the mechanical axis of the dynamometer.[19,24]

With the intention of improving the interpretation of the absolute reliability, the minimum detectable change was assessed, which in terms of practical applications can be used to

Table 3

Mean of test–retest reliability statistics between consecutive testing sessions for the *isokinetic strength variables* for males and females expressed in the original units of measurement (o.u.) with probabilistic inferences presented for the intraclass correlation coefficients (ICC).

Variable (mean)	TE (o.u.) (mean (90% CL))	TE (%)	MDC (o.u.)	MDC (%)	ICC (o.u.) (mean (90% CL))	Inference* ICC
Males						
PT_{EXT}	23.56 (20.96–27.06)	6.26	35.34	9.39	0.91 (0.86–0.95)	Very high
PT_{FLE}	13.04 (9.16–14.44)	5.83	19.56	8.75	0.79 (0.69–0.87)	High
RPT_{EXT}	0.30 (0.27–0.34)	5.95	0.45	8.93	0.90 (0.85–0.94)	High
RPT_{FLE}	0.18 (0.16–0.20)	5.89	0.27	8.84	0.83 (0.75–0.90)	High
TW_{EXT}	191.68 (170.56–220.13)	8.90	287.52	13.35	0.88 (0.82–0.93)	High
TW_{FLE}	86.94 (77.41–100.06)	8.01	130.41	12.02	0.84 (0.76–0.90)	High
RTW_{EXT}	2.59 (2.31–2.98)	9.07	3.89	13.61	0.85 (0.77–0.91)	High
RTW_{FLE}	1.16 (1.03–1.33)	8.02	1.74	12.03	0.84 (0.75–0.90)	High
Females						
PT_{EXT}	19.08 (17.03–21.92)	7.52	28.62	11.28	0.86 (0.79–0.91)	High
PT_{FLE}	12.99 (11.60–14.92)	8.00	19.49	12.00	0.74 (0.63–0.84)	High
RPT_{EXT}	0.32 (0.29–0.37)	7.52	0.48	11.28	0.79 (0.69–0.87)	High
RPT_{FLE}	0.23 (0.20–0.26)	8.23	0.35	12.35	0.81 (0.71–0.88)	High
TW_{EXT}	133.55 (119.21–153.39)	9.12	200.33	13.68	0.87 (0.80–0.92)	High
TW_{FLE}	43.19 (38.53–49.50)	6.46	64.79	9.69	0.87 (0.81–0.92)	High
RTW_{EXT}	2.33 (2.08–2.67)	9.38	3.50	14.07	0.80 (0.70–0.88)	High
RTW_{FLE}	0.73 (0.65–0.83)	6.39	1.10	9.59	0.84 (0.77–0.91)	High

* Terms for ICC magnitudes: low, 0.20–0.50; moderate, 0.50–0.75; high, 0.75–0.90; very high, 0.90–0.99.

Abbreviations: CL = confidence limit; EXT = extension; FLE = flexion; MDC = minimal detectable change; PT = peak torque; RPT = relative peak torque; RTW = relative total work; TE = typical error; TW = total work.

indicate the limit for the smallest change that indicates a real improvement in the measurement after an intervention.[50] The results show that changes over 14% in strength variables and over 20% in endurance variables (Tables 3 and 4) would be needed to ensure that the observed changes are real changes rather than measurement errors or participants' variability. In addition, the general trivial changes observed between consecutive testing sessions for strength and endurance variables may support the idea that no systematic error associated with learning effects occurred.

Table 4

Mean of test–retest reliability statistics between consecutive testing sessions for the *isokinetic endurance variables* for males and females expressed in the original units of measurement (o.u.) with probabilistic inferences presented for the intraclass correlation coefficients (ICC).

Variable	TE (o.u.) (mean (90% CL))	TE (%)	MDC (o.u.)	MDC (%)	ICC (o.u.) (mean (90% CL))	Inference* ICC (o.u.)
Males						
ER_{EXT}	5.06 (4.51–5.83)	5.65	7.59	8.48	0.82 (0.73–0.89)	High
ER_{FLE}	4.38 (3.90–5.05)	5.30	6.57	7.95	0.58 (0.43–0.72)	Moderate
MER_{EXT}	4.10 (3.64–4.70)	5.03	6.15	7.55	0.69 (0.57–0.81)	Moderate
MER_{FLE}	3.85 (3.43–4.45)	4.97	5.78	7.46	0.57 (0.42–0.72)	Moderate
FFR_{EXT}	6.85 (6.09–7.86)	9.52	10.28	14.28	0.68 (0.55–0.80)	Moderate
FFR_{FLE}	7.41 (6.59–8.51)	11.46	11.12	17.19	0.51 (0.35–0.67)	Moderate
RR_{EXT}	10.08 (8.96–11.62)	11.02	15.12	16.53	0.48 (0.33–0.65)	Low
RR_{FLE}	8.46 (7.51–9.77)	10.02	12.69	15.03	0.40 (0.24–0.58)	Low
MRR_{EXT}	7.11 (6.33–8.17)	8.18	10.67	12.27	0.43 (0.27–0.60)	Low
MRR_{FLE}	6.45 (5.75–7.44)	7.80	9.68	11.70	0.52 (0.37–0.68)	Moderate
Females						
ER_{EXT}	4.69 (4.18–5.40)	5.35	7.04	8.03	0.75 (0.64–0.84)	High
ER_{FLE}	6.31 (5.63–7.23)	7.93	9.47	11.90	0.32 (0.17–0.50)	Low
MER_{EXT}	3.97 (3.54–4.55)	4.88	5.96	7.32	0.63 (0.50–0.76)	Moderate
MER_{FLE}	6.43 (5.74–7.37)	8.83	9.65	13.25	0.21 (0.07–0.39)	Low
FFR_{EXT}	6.20 (5.53–7.11)	8.31	9.30	12.47	0.58 (0.44–0.72)	Moderate
FFR_{FLE}	8.03 (7.14–9.21)	13.29	12.05	19.94	0.55 (0.40–0.70)	Moderate
RR_{EXT}	9.55 (8.48–10.96)	10.01	14.33	15.02	0.55 (0.40–0.70)	Moderate
RR_{FLE}	11.09 (9.87–12.71)	13.75	16.64	20.63	0.31 (0.16–0.49)	Low
MRR_{EXT}	5.58 (4.98–6.41)	6.19	8.37	9.29	0.51 (0.36–0.67)	Moderate
MRR_{FLE}	8.79 (7.82–10.07)	11.11	13.19	16.67	0.44 (0.28–0.60)	Low

* Terms for ICC magnitudes: low, 0.20–0.50; moderate, 0.50–0.75; high, 0.75–0.90; very high, 0.90–0.99.

Abbreviations: CL, confidence limit; ER = endurance ratio; EXT = extension; FFR = final fatigue ratio; FLE = flexion; MDC = minimal detectable change; MER = modified endurance ratio; MRR = modified recovery ratio; RR = recovery ratio; TE = typical error.

Interestingly, the reliability obtained in trunk extension exertions was slightly higher than the reliability observed in trunk flexion exertions, mainly in endurance variables. In the present study, the differences between extension and flexion directions could be due to the structure of the dynamometer used. The Biodex isokinetic dynamometer has a rigid support both in the back and in the front of the legs (Fig. 1), helping participants to consistently transmit the forces from the lower extremities to the trunk during extension exertions. However, the dynamometer does not have these rigid structures on the chest or behind the legs, which could make the performance of maximum flexion exertions more difficult and, therefore, less consistent. In this sense, with the goal of enhancing flexion exertion reliability, it would be interesting to modify the dynamometer structure by implementing a rigid support on the chest and behind the legs to allow better force transmission in both phases of the movement.

Regarding reliability differences between males and females, both samples presented similar relative and absolute reliability values. These results support those previously obtained by Delitto et al.[23] but differ from those by Dvir et al.[20] and Keller et al.,[24,25] who found higher reliability values for females and males, respectively. In addition, most isokinetic studies of other muscle groups have shown worse reliability results in males than in females, probably as a result of a higher difficulty of controlling the males' body position during the protocol. In this sense, the higher anthropometric dimensions and the higher experience in maximum efforts of some males in these studies may have allowed them to exert higher forces,[25] so inappropriate strapping could have changed the initial position, affecting the pelvic axis alignment.[19,24,40] On the contrary, the lack of reliability differences between sexes in the current study could be caused by (1) an adequate position standardization in the attachment of the dynamometer via the different adjustable pads and straps used; (2) the control of body position across the sets; and/or (3) similar male and female experience with performing maximum efforts, which could decrease the differences between them. Although males and females obtained similar reliability values, both samples showed large differences in trunk muscle performance (Tables 1 and 2). Males showed higher trunk flexion and extension strength and higher trunk flexion endurance than females, but no sex differences were observed for trunk extension endurance.

For a comprehensive analysis of isokinetic protocol reliability, the learning effect was assessed through 5 testing sessions. Although a few significant differences were found for total work and relative total work, generally the strength and endurance variables showed no significant differences between sessions in both sexes (Tables 1 and 2), demonstrating the consistency of the measurements. In addition, when small differences were found for the strength variables, a reduction between the first and the rest of the sessions was observed (Fig. 3), which cannot be interpreted as a learning effect of the protocol. The reason for this decrease may be the lack of motivation of the participants because of the extensive and intensive demands of the protocol (i.e., 4×15 maximum flexion–extension exertions) and the long study duration (i.e., 5 testing sessions in 8 weeks). Therefore, only 1 session would be enough to obtain reliable strength and endurance values in this protocol. These results are supported by previous studies that also analyzed strength variables and found no significant differences between sessions.[18–21,23–26]

Application of the data of this study is limited to healthy and physically active young males and females. Future investigations should include individuals with different spinal conditions, ages, physical activity levels, and so on. In this sense, because of the high physical demands of this protocol, some modifications may be needed for untrained individuals or those with low back injury (e.g., increasing warm-up duration, reducing angular velocity and number of sets, *etc.*). In addition, as has been explained earlier, our results are influenced by the characteristics of the dynamometer used in this study (e.g., adjustable pads and straps, rigid supports, *etc.*). Thus, if this protocol is carried out using other dynamometers, it would be advisable to perform new reliability analyses. Another limitation of this study is that the participants' body mass was measured only in the first testing session. Although researchers did not appreciate significant weight variations in the participants, and the reliability of the relative peak torque and relative total work (variables that depend on a participant's body mass) was high, anthropometry changes throughout the study could affect our results.

5. Conclusion

The findings of this study provide trainers and researchers with a 10-min single-session protocol to perform a reliable muscle strength and endurance evaluation of trunk flexor and extensor muscles, all within the same protocol. Based on the good reliability results obtained for all strength variables, any of them could be used to assess trunk muscle strength in physically active young males and females. However, regarding the endurance variables, endurance ratio and modified endurance ratio showed the best reliability results, mainly in the extension direction and in males.

References

1. Barbado D, Lopez-Valenciano A, Juan-Recio C, Montero-Carretero C, van Dieën JH, Vera-Garcia FJ. Trunk stability, trunk strength and sport performance level in judo. *PLoS ONE* 2016;**11**:e0156267.
2. Chan RH. Endurance times of trunk muscles in male intercollegiate rowers in Hong Kong. *Arch Phys Med Rehabil* 2005;**86**:2009–12.
3. Correia JP, Oliveira R, Vaz JR, Silva L, Pezarat-Correia P. Trunk muscle activation, fatigue and low back pain in tennis players. *J Sci Med Sport* 2016;**19**:311–16.

4. Evans K, Refshauge KM, Adams R, Aliprandi L. Predictors of low back pain in young elite golfers: a preliminary study. *Phys Ther Sport* 2005;**6**:122–30.

5. Young J. Trunk contributions to baseball pitching velocity. *Undergraduate senior thesis*. Columbus, OH: The Ohio State University, 2014.

6. Behm DG, Wahl MJ, Button DC, Power KE, Anderson KG. Relationship between hockey skating speed and selected performance measures. *J Strength Cond Res* 2005;**19**:326–31.

7. Hibbs AE, Thompson KG, French D, Wrigley A, Spears I. Optimizing performance by improving core stability and core strength. *Sports Med* 2008;**38**:995–1008.

8. Clayton MA, Trudo CE, Laubach LL, Linderman JK, De Marco GM, Barr S. Relationships between isokinetic core strength and field based athletic performance tests in male collegiate baseball players. *J Exerc Physiol Online* 2011;**14**:20–30.

9. Prieske O, Muehlbauer T, Borde R, Gube M, Bruhn S, Behm D, et al. Neuromuscular and athletic performance following core strength training in elite youth soccer: role of instability. *Scand J Med Sci Sports* 2016;**26**:48–56.

10. Durall CJ, Udermann BE, Johansen DR, Gibson B, Reineke DM, Reuteman P. The effects of preseason trunk muscle training on low-back pain occurrence in women collegiate gymnasts. *J Strength Cond Res* 2009;**23**:86–92.

11. Williams CA, Singh M. Dynamic trunk strength of Canadian football players, soccer players, and middle to long distance runners. *J Orthop Sports Phys Ther* 1997;**25**:271–6.

12. Iwai K, Okada T, Nakazato K, Fujimoto H, Yamamoto Y, Nakajima H. Sport-specific characteristics of trunk muscles in collegiate wrestlers and judokas. *J Strength Cond Res* 2008;**22**:350–8.

13. Lee JH, Hoshino Y, Nakamura K, Kariya Y, Saita K, Ito K. Trunk muscle weakness as a risk factor for low back pain: a 5-year prospective study. *Spine* 1999;**24**:54–7.

14. Yahia A, Ghroubi S, Kharrat O, Jribi S, Elleuch M, Elleuch M. A study of isokinetic trunk and knee muscle strength in patients with chronic sciatica. *Ann Phys Rehabil Med* 2010;**53**:239–49.

15. Bayramoglu M, Akman MN, Klnç S, Çetin N, Yavuz N Özker R. Isokinetic measurement of trunk muscle strength in women with chronic low-back pain. *Am J Phys Med Rehabil* 2001;**80**:650–5.

16. Ganzit G, Chisotti L, Albertini G, Martore M, Gribaudo C. Isokinetic testing of flexor and extensor muscles in athletes suffering from low back pain. *J Sports Med Phys Fit* 1998;**38**:330–6.

17. Drouin JM, Valovich-mcLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex System 3 Pro isokinetic dynamometer velocity, torque and position measurements. *Eur J Appl Physiol* 2004;**91**:22–9.

18. Cowley PM, Fitzgerald S, Sottung K, Swensen T. Age, weight, and the front abdominal power test as predictors of isokinetic trunk strength and work in young men and women. *J Strength Cond Res* 2009;**23**:915–25.

19. Dervisevica E, Hadzic V, Burger H. Reproducibility of trunk isokinetic strength findings in healthy individuals. *Isokinet Exerc Sci* 2007;**15**:99–109.

20. Laughlin MS, Lee SM, Loehr JA, Amonette WE. *Isokinetic strength and endurance tests used pre-and post-spaceflight: test-retest reliability*. NASA/TM-2009-214787. Hanover, MD: NASA STI Program Office, 2009.

21. Müller J, Müller S, Stoll J, Fröhlich K, Baur H, Mayer F. Reproducibility of maximum isokinetic trunk strength testing in healthy adolescent athletes. *Sports Orthop Traumatol* 2014;**30**:229–37.

22. Gómez HB, Zarco RC, Arias DC, García MPD, Hernández SRL, Martínez E. Valoración isocinética del tronco en sujetos asintomáticos del Centro Nacional de Rehabilitación. *Acta Ortop Mex* 2005;**19**:49–55.

23. Delitto A, Rose SJ, Crandell CE, Strube MJ. Reliability of isokinetic measurements of trunk muscle performance. *Spine* 1991;**16**:800–3.

24. Dvir Z, Keating J. Reproducibility and validity of a new test protocol for measuring isokinetic trunk extension strength. *Clin Biomech (Bristol, Avon)* 2001;**16**:627–30.

25. Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sørensen test, and Åstrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. *Spine* 2001;**26**:771–7.

26. Wessel J, Ford D Van Driesum D. Measurement of torque of trunk flexors at different velocities. *Scand J Rehabil Med* 1992;**24**:175–80.

27. Biering-Sørensen F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine* 1984;**9**:106–19.

28. Brotons-Gil E, García-Vaquero MP, Peco-González N, Vera-Garcia FJ. Flexion-rotation trunk test to assess abdominal muscle endurance: reliability, learning effect, and sex differences. *J Strength Cond Res* 2013;**27**:1602–8.

29. Knudson D. Validity and reliability of a bench trunk-curl test of abdominal endurance. *J Strength Cond Res* 2001;**15**:81–5.

30. McGill SM, Childs A, Liebenson C. Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database. *Arch Phys Med Rehabil* 1999;**80**:941–4.

31. Linderman JK, Board R, Astorino T, Lee WL. The relationship between core strength and performance in division I female soccer players. *J Exerc Physiol Online* 2009;**12**:21–8.

32. Recio CJ, Murillo DB, Valenciano AL, Vera-García FJ. Test de campo para valorar la resistencia de los músculos del tronco. *Apunts Educ Fís Deportes* 2014;**59–68**.

33. Evans K, Refshauge KM, Adams R. Trunk muscle endurance tests: reliability, and gender differences in athletes. *J Sci Med Sport* 2007;**10**:447–55.

34. Moreland J, Finch E, Stratford P, Balsor B, Gill C. Interrater reliability of six tests of trunk muscle function and endurance. *J Orthop Sports Phys Ther* 1997;**26**:200–8.

35. Kraemer WJ, Ratamess NA. Fundamentals of resistance training: progression and exercise prescription. *Med Sci Sports Exerc* 2004;**36**:674–88.

36. Lindsay DM, Horton JF. Trunk rotation strength and endurance in healthy normals and elite male golfers with and without low back pain. *N Am J Sports Phys Ther* 2006;**1**:80–9.

37. Mayer T, Gatchel R, Betancur J, Bovasso E. Trunk muscle endurance measurement: isometric contrasted to isokinetic testing in normal subjects. *Spine* 1995;**20**:920–6.

38. Grabiner M, Jeziorowski J. Isokinetic trunk extension and flexion strength-endurance relationships. *Clin Biomech (Bristol, Avon)* 1991;**6**:118–22.

39. Watkins M, Harris B. Evaluation of isokinetic muscle performance. *Clin Sports Med* 1983;**2**:37–53.

40. Karatas GK, Gögüs F, Meray J. Reliability of isokinetic trunk muscle strength measurement. *Am J Phys Med Rehabil* 2002;**81**:79–85.

41. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;**30**:1–15.

42. Juan-Recio C, Lopez-Vivancos A, Moya M, Sarabia J, Vera-Garcia F. Short-term effect of crunch exercise frequency on abdominal muscle endurance. *J Sports Med Phys Fit* 2015;**55**:280–9.

43. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;**26**:217–38.

44. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.

45. Hopkins W, Marshall S, Batterham A, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009;**41**:3–13.

46. Munro BH. *Statistical methods for health care research*. Philadelphia: Lippincott Williams & Wilkins; 2005.

47. Denegar CR, Ball DW. Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement. *J Sport Rehab* 1993;**2**:35–42.

48. Newton M, Waddell G. Trunk strength testing with iso-machines: part 1: review of a decade of scientific evidence. *Spine* 1993;**18**:801–11.

49. Vincent W, Weir J. *Statistics in kinesiology*. Champaign, IL: Human Kinetics; 1994.

50. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005;**19**:231–40.