Review Paper DOI: https://doi.org/10.5114/biolsport.2022.106149

Test-retest reliability of isometric mid-thigh pull maximum strength assessment: a systematic review

AUTHORS: Jozo Grgic[1], Bela Scapec[2], Pavle Mikulic[2], Zeljko Pedisic[1]

[1] Institute for Health and Sport, Victoria University, Melbourne, Australia
[2] Faculty of Kinesiology, University of Zagreb, Zagreb, Croatia

ABSTRACT: The aim of this systematic review was to explore the test-retest reliability of isometric mid-thigh pull maximum strength assessment. We searched through five databases to find studies that examined the test-retest reliability of peak force in the isometric mid-thigh pull exercise. From each included study, we extracted intra-class correlation coefficients (ICC) and/or coefficient of variation (CV). The methodological quality of the included studies was evaluated using the COSMIN checklist. A total of 16 good-to-excellent quality studies were included in the review. When considering results from all included studies, ICCs ranged from 0.73 to 0.99 (median ICC = 0.96), where 78% of ICCs were ≥ 0.90, and 98% of ICCs were ≥ 0.75. The range of reported CVs was from 0.7% to 11.1% (median CV = 4.9%), where 58% of CVs were ≤ 5%. Reliability was also good-to-excellent for both relative and absolute peak force and for both bilateral and unilateral isometric mid-thigh pull tests. The majority of studies did not find significant differences between testing sessions. It can be concluded that the isometric mid-thigh pull maximum strength assessment has good-to-excellent test-retest reliability. The isometric mid-thigh pull maximum strength assessment can be used as a reliable test in sports practice and for research purposes.

CITATION: Grgic J, Scapec B, Mikulic P, Pedisic Z. Test-retest reliability of isometric mid-thigh pull maximum strength assessment: a systematic review. Biol Sport. 2022;39(2):407–414.

Received: 2021-01-08; Reviewed: 2021-02-07; Re-submitted: 2021-02-10; Accepted: 2021-03-20; Published: 2021-06-01.

Corresponding author:
Jozo Grgic
Institute for Health and Sport (IHES), Victoria University, Melbourne, Australia
Email: jozo.grgic@live.vu.edu.au

ORCID:
Jozo Grgic
0000-0002-6929-2844

Pavle Mikulic
0000-0003-1244-631X

Zeljko Pedisic
0000-0003-2886-3556

Key words:
Reliability
Repeatability
Reproducibility
Testing
Strength

INTRODUCTION

Muscular strength can be defined as "the ability to exert force on an external object or resistance" [1]. Maximum muscular force production is influenced by muscle activation, size, and architecture [2]. There are several tests of maximal strength, and one of them is the isometric mid-thigh pull test. The isometric mid-thigh pull test was developed in the 1990s, and it has gained substantial popularity in recent years [3]. As suggested by its name, the test involves maximum isometric force production, while holding a barbell that is set in the mid-thigh position [3, 4]. This test was originally developed to mimic the second pull phase of the clean, making it highly suitable to weightlifting [3, 4]. However, the test has also been used for strength assessment in other sports and for research purposes [3, 4]. Important advantages of this test is that it is not overly fatiguing and that it generally takes less time than one-repetition maximum (1RM) assessment [3]. Furthermore, the isometric mid-thigh pull strength is considered safe, with a very low reported incidence of injury [3]. Unlike single-joint isometric tests of strength, the isometric mid-thigh pull test is generally strongly associated with dynamic exercise performance [4, 5, 6]. For example, peak force recorded in this test is in correlation with: 1RM in the clean-and-jerk, snatch, squat, and deadlift (r = 0.64–0.97) [7, 8, 9]; change of direction performance (r = 0.57–0.85) [10, 11]; sprinting kinetics (r = 0.48–0.73) [12]; and vertical jump height (r = 0.59–0.82) [13, 14].

This test has been used in studies that evaluated the effects of resistance training programs on maximum strength development, the acute ergogenic effects of supplements on strength (e.g., sodium bicarbonate, caffeine), and the influence of motor learning strategies on maximum force production [15, 16, 17, 18]. Peak force from the isometric mid-thigh pull test is also one of the components of the "Dynamic Strength Index" (i.e., the ratio of ballistic peak force and isometric peak force) [19]. This index is commonly used to develop training programs and to evaluate their efficacy [19]. Additionally, the isometric mid-thigh pull test is also used to assess fatigue and recovery from exercise and competition [20, 21]. This test has also been utilized to evaluate physical characteristics of academy rugby union players and to track changes in maximal strength of baseball players across competitive seasons [22, 23]. The application of the isometric mid-thigh pull test in these various areas of research demonstrates its growing popularity in recent years.

Given the increased use of this test in research and applied settings, it is important to establish its test-retest reliability. Test-retest reliability refers to the consistency of results in a given test across repeated measures [24, 25]. Poor test-retest reliability may increase the probability of type II error [24, 25]. While several studies explored the test-retest reliability of maximum strength using the isometric mid-thigh pull test, the findings varied across the studies [9, 19, 21, 26–38]. Therefore, the aim of this paper was to conduct a systematic review of studies that explored the test-retest reliability of isometric mid-thigh pull maximum strength assessment, and to summarize their findings.

MATERIALS AND METHODS

Search strategy

For this review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [39]. The following search syntax was used to identify the studies that examined the test-retest reliability of maximum strength assessment using the isometric mid-thigh pull test: ("mid-thigh pull" OR "mid thigh pull" OR "midthigh pull" OR "mid-thigh clean" OR "mid thigh clean" OR "midthigh clean") AND (reliability OR repeatability OR reproducibility). We searched through five databases, including: CINAHL (through EBSCOhost), PubMed/MEDLINE, Scopus, SPORTDiscus (through EBSCOhost), and Web of Science (including all Citation Indexes). In addition to the primary search, we performed a secondary search that consisted of examining reference lists of all included studies.

Inclusion criteria

We included studies that satisfied the following criteria: (a) published in a peer-reviewed journal and in English language; (b) examined the test-retest reliability of relative or absolute peak force in the unilateral or bilateral isometric mid-thigh pull exercise; and (c) presented intra-class correlation coefficient (ICC) and/or coefficient of variation (CV). The search and study selection according to the eligibility criteria were concluded on December 1st, 2020. They were performed independently by two authors of the review (JG and BS), to prevent possible study selection bias.

Data extraction

From each of the included studies, we extracted the following data: (a) participant characteristics; (b) time between testing sessions; (c) familiarization with the test and warm-up protocol; (d) hip and knee angle used for the test; and (e) ICC and/or CV values for peak force. Data extraction was performed independently by two authors of the review (JG and BS). Following the completion of the data extraction by both authors, the files with extracted data were compared, discussed, and harmonized.

Reliability data interpretation

ICCs were interpreted using the thresholds proposed by Koo and Li [40]. Specifically, we classified ICCs as indication of: "poor reliability" (ICC < 0.50), "moderate reliability" (ICC $= 0.50$–0.75), "good reliability" (ICC $= 0.76$–0.90), or "excellent reliability" (ICC > 0.90). Even though there are no universally accepted thresholds for interpreting CV, in the medical and health research area, CVs of $\leq 5\%$ are generally considered to indicate excellent reliability [41].

Methodological quality

We evaluated the methodological quality of included studies using the form B of *COnsensus-based Standards for the selection of health Measurement Instruments* (COSMIN) checklist [42]. Form B of this checklist has a total of 11 items. These items refer to the number of testing sessions, the days between sessions, test administration, data reporting, methodological limitations, and the sample size adequacy. A study may be given one point per item. Thus, the maximum score on the checklist is 11 points. Based on the summary scores, we classified studies as "excellent quality", "moderate quality", and "poor quality", if they scored 9–11 points, 6–8 points, and fewer than 6 points, respectively [43]. Two authors of the review (JG and BS) independently performed the methodological quality evaluation. Their independent assessments were compared, discussed, and harmonized.

RESULTS

Search results

A total of 721 references were assessed against the eligibility criteria. Specifically, the primary search yielded a total of 190 references (Figure 1). After excluding ineligible studies based on title, abstract, or full-text, 15 studies [9, 19, 21, 26–34, 36–38] were included in the review. In the secondary search, 531 references were assessed and one additional study was included in the review [35]. Therefore, the final number of included studies was 16 [9, 19, 21, 26–38].

Study characteristics

Sample sizes in the included studies ranged from 8 to 59 participants (median = 16). The overall number of participants across all included studies was 307. Thirteen studies included athletes from various sports (e.g., rugby, soccer, netball); two studies were conducted among resistance-trained individuals, and one study included recreationally active males (Table 1). Fifteen studies presented ICCs, while 14 studies presented CVs. The time between testing sessions ranged from 1 to 7 days. The warm-up protocol varied across studies, but it most commonly included two practice attempts before the test, performed at 50% and 75% of perceived maximum effort.

Methodological quality

Fifteen studies were classified as being of excellent methodological quality, while one study was classified as having moderate methodological quality (Table 2). The scores for individual studies on all items of the COSMIN checklist are presented in Table 2.

Overall reliability

When considering results from all included studies, ICCs ranged from 0.73 to 0.99 (median ICC $= 0.96$), where 78% of ICCs were ≥ 0.90, and 98% of ICCs were ≥ 0.75. The range of reported CVs was from 0.7% to 11.1% (median CV $= 4.9\%$), where 58% of CVs were $\leq 5\%$.

TABLE 1. Summary of included studies

Study	Sample	Time between testing sessions	Familiarization	Warm-up protocol	Sampling rate	Hip and knee angle	ICC (95% CI)	CV (95% CI)
Aben et al. [21]	10 male rugby league players	7 days	3 practice sessions	Dynamic warm-up including lunges, leg swings and other exercises	1000 Hz	Knee: 120–135° Hip: 140–150°	Peak force: 0.92 (0.68, 0.98)	Peak force: 4.3% (3.2%, 7.2%)
Comfort et al. [26]	8 male college athletes	7 days	1 practice session	1 set of 3 repetitions of mid-thigh power cleans at 40%, 60%, and 80% of 1RM power clean	600 Hz	120° knee, 125° hip 130° knee, 125° hip 140° knee, 125° hip 150° knee, 125° hip 120° knee, 145° hip 130° knee, 145° hip 140° knee, 145° hip 150° knee, 145° hip Preferred position	*120° knee, 125° hip* Peak force: 0.99 *130° knee, 125° hip* Peak force: 0.99 *140° knee, 125° hip* Peak force: 0.99 *150° knee, 125° hip* Peak force: 0.99 *120° knee, 145° hip* Peak force: 0.99 *130° knee, 145° hip* Peak force: 0.99 *140° knee, 145° hip* Peak force: 0.98 *150° knee, 145° hip* Peak force: 0.99 *Preferred position* Peak force: 0.99	Not reported
Comfort et al. [27]	29 male collegiate athletes	3 days	Prior experience with the exercise	3 test attempts at 50%, 75%, and 90% of perceived maximum effort	1000 Hz	Knee: 139.5 ± 3.3° Hip: 145.1 ± 3.4°	Peak force: 0.98 (0.95, 0.99)	Peak force: 0.7%
De Witt et al. [9]	9 resistance-trained participants (5 males and 4 females)	At least 3 days	1 practice session	2 test attempts at 50% and 75% of perceived maximum effort	1000 Hz	Knee: flexed at 36 ± 3° Hip: flexed at 43 ± 3°	0.89 (0.74, 1.00)	Not reported
Dos'Santos et al. [28]	13 male youth soccer players	2 days	Prior experience with the exercise	5 minutes of dynamic stretching, 1 set of 5 repetitions of mid-thigh clean pulls, and 2 test attempts at 50% and 75% of perceived maximum effort	1000 Hz	Knee: 137–146° Hip: 140–149°	Peak force: 0.96 (0.88, 0.99)	Peak force: 4.6% (3.3%, 7.7%)
Dos'Santos et al. [29]	10 collegiate athletes	7 days	Prior experience with the exercise	10 bodyweight squats and lunges and 2 test attempts at 50%, and 75% of perceived maximum effort	1000 Hz	145° knee, 145° hip 145° knee, 175° hip	*145° knee, 145° hip* Peak force: 0.97 (0.81, 0.99) *145° knee, 175° hip* Peak force: 0.97 (0.89, 0.99)	*145° knee, 145° hip* Peak force: 4.5% (1.6%, 7.3%) *145° knee, 175° hip* Peak force: 5.3% (3.0%, 7.6%)
Guppy et al. [30]	17 strength and power athletes (11 males and 6 females)	7 days	1 practice session	1 set of 3 repetitions of dynamic mid-thigh pulls at 40%, 60% and 80% of 1RM clean; and 2 test attempts at 50% and 75% of perceived maximum effort	1000 Hz	145° knee, 145° hip, and traditional barbell position 145° knee, 120° hip, and traditional barbell position 120° knee, 125° hip, and mid-thigh barbell position 120° knee, 145° hip, and mid-thigh barbell position	*145° knee, 145° hip, and traditional barbell position* Peak force: 0.98 (0.95, 0.99) *145° knee, 120° hip, and traditional barbell position* Peak force: 0.97 (0.91, 0.99) *120° knee, 125° hip, and mid-thigh barbell position* Peak force: 0.84 (0.59, 0.94) *120° knee, 145° hip, and mid-thigh barbell position* Peak force: 0.92 (0.78, 0.97)	*145° knee, 145° hip, and traditional barbell position* Peak force: 4.0% (2.9%, 6.2%) *145° knee, 120° hip, and traditional barbell position* Peak force: 5.0% (3.7%, 7.8%) *120° knee, 125° hip, and mid-thigh barbell position* Peak force: 11.1% (8.1%, 17.7%) *120° knee, 145° hip, and mid-thigh barbell position* Peak force: 8.0% (5.9%, 12.7%)
Haines et al. [31]	17 male adolescent athletes	2 consecutive days	8 practice sessions	2 sub-maximal test attempts	500 Hz	Knee: 145–150° Hip: not reported	Peak force: 0.87 (0.71, 0.95) Peak force (relative): 0.73 (0.45, 0.88)	Peak force: 6.4% (4.9%, 9.4%) Peak force (relative): 6.4% (5.0%, 9.4%)

TABLE 1. Continue

Study	Sample	Time between testing sessions	Familiarization	Warm-up protocol	Sampling rate	Hip and knee angle	ICC (95% CI)	CV (95% CI)
James et al. [32]	15 recreationally active males	2–7 days	Practice before the main attempts	3 sets of 5 repetitions of dynamic mid-thigh clean pulls using 30–50% of 1RM	1000 Hz	Knee: 141.9 ± 4.3° Hip: 139.2 ± 4.1°	Peak force: 0.96 (0.90, 0.98)	Peak force: 3.1% (2.4%, 4.6%)
Moeskops et al. [33]	19 pre-PHV and 19 post-PHV female athletes	At least 1 day	1 practice session	10 minute dynamic warm-up	1000 Hz	Knee: 140 ± 5° Hip: 135 ± 5°	*Pre-PHV athletes* Peak force: 0.95 (0.83, 0.98) Peak force (relative): 0.81 (0.58, 0.92) *Post-PHV athletes* Peak force: 0.92 (0.80, 0.97) Peak force (relative): 0.81 (0.57, 0.92)	*Pre-PHV athletes* Peak force: 10.2% (7.6%, 15.5%) Peak force (relative): 10.1% (7.5%, 15.3%) *Post-PHV athletes* Peak force: 6.7% (5.0%, 10.0%) Peak force (relative): 7.3% (5.5%, 11.0%)
Sawczuk et al. [34]	59 youth sport athletes (39 males and 20 females)	7 days	Prior experience with the exercise	2 test attempts at 50% and 75% of perceived maximum	Not reported	Knee: 120–135° Hip: not reported	Not reported	5.5% (4.5%, 6.9%)
Sheppard et al. [35]	18 athletes (15 males and 3 females)	2 days	1 practice session	Skipping for 5-min, and 2 sets of 10 bodyweight squats	200 Hz	Knee: 130° Hip: 155–165°	Peak force: 0.99	Peak force: 2.0%
Suarez et al. [36]	13 resistance-trained men	7 days	Prior experience with the exercise and 1 practice session	2 test attempts at 50% and 75% of perceived maximum	1000 Hz	Knee: 120–130°	*Short duration protocol* Peak force: 0.95 (0.85, 0.99) *Long duration protocol* Peak force: 0.96 (0.89, 0.99)	*Short duration protocol* Peak force: 5.1% (3.7%, 8.4%) *Long duration protocol* Peak force: 4.1% (2.9%, 6.8%)
Thomas et al. [19]	19 male college athletes	2 days	2 practice sessions	2 test attempts at 50% and 75% of perceived maximum	600 Hz	Self-selected knee and hip angles	Peak force: 0.95	Peak force: 3.8%
Thomas et al. [37]	17 adolescent athletes (8 males and 9 females)	7 days	1 practice session	2 test attempts at 50% and 75% of perceived maximum	1000 Hz	Self-selected knee and hip angles	*Bilateral* Peak force: 0.86 Peak force (relative): 0.86 *Single leg (left)* Peak force: 0.94 Peak force (relative): 0.89 *Single leg (right)* Peak force: 0.96 Peak force (relative): 0.94	*Bilateral* Peak force: 6.8% Peak force (relative): 6.8% *Single leg (left)* Peak force: 4.0% Peak force (relative): 4.0% *Single leg (right)* Peak force: 3.4% Peak force (relative): 3.4%
Thomas et al. [38]	16 female netball athletes	7 days	Prior experience with the exercise	2 test attempts at 50% and 75% of perceived maximum	600 Hz	Knee: 130–150° Hip: 140–160°	*Single leg (left)* Peak force: 0.95 (0.89, 0.98) Peak force (relative): 0.92 (0.82, 0.97) *Single leg (right)* Peak force: 0.97 (0.93, 0.99) Peak force (relative): 0.94 (0.87, 0.98)	*Single leg (left)* Peak force: 4.9% (3.8%, 7.1%) Peak force (relative): 4.9% (3.8%, 7.1%) *Single leg (right)* Peak force: 4.2% (3.2%, 6.0%) Peak force (relative): 4.2% (3.2%, 6.0%)

PHV: peak height velocity; ICC: intraclass correlation coefficient; CV: coefficient of variation; CI: confidence interval; 1RM: one-repetition maximum

TABLE 2. Methodological quality assessment of the included studies using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	TS
Aben et al. [21]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Comfort et al. [26]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Comfort et al. [27]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
De Witt et al. [9]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Dos'Santos et al. [28]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Dos'Santos et al. [29]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Guppy et al. [30]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Haines et al. [31]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
James et al. [32]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Moeskops et al. [33]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Sawczuk et al. [34]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	No	8
Sheppard et al. [35]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Suarez et al. [36]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Thomas et al. [19]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Thomas et al. [37]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9
Thomas et al. [38]	Yes	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	No	Yes	9

TS: total score

Reliability of bilateral isometric mid-thigh pull

Fifteen studies explored the reliability of peak force in the bilateral isometric mid-thigh pull. The respective ICCs ranged from 0.73 to 0.99 (median ICC = 0.96), where 75% of ICCs were ≥ 0.90, and 97% of ICCs were ≥ 0.75. The range of reported CVs was from 0.7% to 11.1% (median CV = 5.3%), where 43% of CVs were ≤ 5%.

Reliability of unilateral isometric mid-thigh pull

Two studies explored the reliability of peak force in the unilateral isometric mid-thigh pull. They reported separate ICCs for both limbs. The ICCs ranged from 0.89 to 0.97 (median ICC = 0.94), where 88% of ICCs were ≥ 0.90. The range of reported CVs was from 3.4% to 4.9% (median CV = 4.1%), where all of CVs were ≤ 5%.

Reliability of absolute peak force

Sixteen studies explored the reliability of absolute peak force. The respective ICCs ranged from 0.84 to 0.99 (median ICC = 0.97), where 88% of ICCs were ≥ 0.90. The range of reported CVs was from 0.7% to 11.1% (median CV = 4.6%), where 61% of CVs were ≤ 5%.

Reliability of relative peak force

Four studies explored the reliability of relative peak force. The respective ICCs ranged from 0.73 to 0.94 (median ICC = 0.88), where 38% of ICCs were ≥ 0.90, and 88% of ICCs were ≥ 0.75. The range of reported CVs was from 3.4% to 10.1% (median CV = 5.7%), where 50% of CVs were ≤ 5%.

Systematic changes in results between repeated measurements

Ten studies [21, 27–33, 36, 37] examined if there were any systematic changes in the results between the test and retest sessions. Eight studies did not find significant differences between the testing sessions. One study [29] found higher peak force values in the first testing session, and one study [33] found higher values in the second testing session.

DISCUSSION

The main finding of this review is that peak force assessment in the isometric mid-thigh pull exercise has good-to-excellent test-retest reliability. The reliability is good-to-excellent for both absolute and relative peak force and for both bilateral and unilateral isometric mid-thigh pull tests. The majority of studies did not find significant differences in peak force between testing sessions. From a practical perspective, we conclude that sport and exercise practitioners can use the bilateral and unilateral isometric mid-thigh pull tests as reliable measures of relative and absolute peak force. Additionally, practitioners should also consider that peak force values obtained in this test do not seem to change systematically between repeated measures.

The position of knees and hips during the isometric mid-thigh pull test is an important methodological aspect often discussed in previous research [4]. One study evaluated peak force values produced during this test when using a "bent" position (125° knee and 125° hip angles) and an "upright" position" (125° knee and 145° hip angles) [44]. Peak force was substantially higher when using the "upright" position [44]. This variation in peak force production indicates that knee and hip angles are highly relevant for this test. However, it seems that the test-retest reliability of peak force is generally good-to-excellent, regardless of the specific angles used, as long as they are standardized. This has been thoroughly explored by Comfort et al. [26], who evaluated reliability while using nine different knee and hip angle combinations. Knee and hip angles in their study ranged 120°-150° and 125°-145°, respectively. Additionally, the participants in their study were allowed to self-select their preferred position for the test. All ICCs for peak force from this study were very high, ranging from 0.98 to 0.99, which indicates that reliability is not significantly affected by changes in knee and hip angles. Another study [29] used two different positions (145° knee and 145° hips vs. 145° knee and 175° hips) and also observed very high reliability (ICC: 0.97; CV: 4.5% to 5.1%) for both variants of the test. Therefore, while the reliability of peak force does not seem to be affected by different hip and knee angle combinations, practitioners should consider that optimal results in the isometric mid-thigh pull are obtained when using a knee angle of 125–145° and a hip angle of 140–150°, respectively [4].

When initially developed, the isometric mid-thigh pull was performed bilaterally [3, 4]. However, this test variation does not provide data on the isolated limb's force production capability. Therefore, researchers have designed a unilateral version of the isometric mid-thigh pull test. Assessing muscle asymmetry by using the unilateral version of this test may be of substantial practical importance, given that lower limb strength imbalance may be associated with a higher risk of injury [45]. Additionally, some common movements in sport (e.g., sprinting and change of direction) are also unilateral. Still, it should be considered that factors such as balance during the test may impact performance and, subsequently, the reliability of the unilateral version of this test. While only two studies [37, 38] explored the test-retest reliability of peak force using the unilateral isometric mid-thigh pull, both of them found good-to-excellent reliability. However, given that the populations in these two studies included adolescent athletes and netball players, more research on this topic is needed. Future studies should also explore the reliability of this test among older adults, given that a lack of strength symmetry may increase gait variability, gait asymmetry, and risk of falls in this population, making the unilateral isometric mid-thigh pull test a potentially very useful diagnostic tool [46].

Current recommendations for exercise testing are that participants should be familiarized with the testing protocol [25]. Out of the 16 included studies, 10 incorporated at least one familiarization session as a part of their study designs (Table 1). Five studies did not use a familiarization session, but they included participants who had prior experience with this specific test (Table 1). Similar reliability values were observed across all studies, which would suggest that familiarization with this test might not be of major importance, as long as participants have prior experience with the test. Indeed, one study [32] included recreationally active participants and did not use a familiarization session, as the authors only allowed two practice attempts with the mid-thigh pull exercise as a part of the warm-up. Despite the lack of familiarization session, this study reported an ICC of 0.96 and a CV of 3.1%, indicating a high reliability of the test. These results suggest that familiarization with the test may be incorporated in the warm-up exercise before the main attempts. Such approach may save time. However, optimal practice is to include at least one separate familiarization session [4].

Some maximum strength tests are somewhat affected by systematic changes in results between repeated measurements. For example, a recent review of the reliability of the 1RM test found that 34% of 32 included studies reported systematic changes between the test and retest sessions [43]. Specifically, with the 1RM test, there seems to be a small learning effect, given that the participants tend to perform better in the second testing session. Ten studies included in the current review examined potential differences in performance between the testing sessions, and only two found significant changes. In one study [29], peak force was higher in the first session, while in the other study, peak force was higher in the second session [33]. Regardless of the direction, the differences between testing sessions were small (Cohen's *d*: 0.21 to 0.29). Additionally, if there is any learning effect for peak force in the isometric mid-thigh pull test, it might be age-dependent. One study included a group of pre-peak height velocity (age: 8.0 ± 2.0 years) and a group of post-peak height velocity (age: 14.6 ± 1.5 years) female athletes [33]. Systematic changes between the test and retest sessions were observed only in the pre-peak height velocity group. To enable drawing sound conclusions about possible age-dependency of the learning effect, these findings have to be confirmed in future studies.

Limitations

In this review we focused primarily on the test-retest reliability of peak force in the isometric mid-thigh pull. We opted to focus specifically on this variable, because it is commonly found to be highly correlated with various aspects of dynamic exercise performance [4]. However, focusing only on peak force might be considered as a limitation of our review, given that several other biomechanical variables, such as impulse, rate of force development, and force at specific time points can be obtained from isometric mid-thigh pull testing [4]. Information about the reliability of these variables can be found in the narrative review by Brady et al. [47].

There are different types of ICC, each designed for specific purposes [40]. When calculated from the same data, six different types of ICC ranged from 0.51 to 0.87 [48]. While the selection of an adequate type of ICC is undoubtedly important, only two included studies [21, 27] explicitly stated which type of ICC they used, which is another limitation of this review that should be considered. Future primary studies should specify the ICC type used in the analysis of reliability, to allow for better-informed comparisons of results between studies.

Methodological quality

The included studies were classified as being of good-to-excellent methodological quality, based on the COSMIN checklist. Still, it should also be considered that none of the studies satisfied item 3 on the COSMIN checklist, which refers to the adequacy of sample size. According to the COSMIN checklist, at least 100 participants should be included in studies on test-retest reliability [42]. However, this threshold can be considered arbitrary, given that the required sample size will, in addition to the number of measurements, also depend on the expected ICC, significance level, and the acceptable width of the confidence interval. For example, according to the equation provided by Bonett [49], for an expected ICC of 0.95 with a width of the 95% confidence interval of ± 0.03 the required sample size is 42 participants, while for the width of the 95% confidence interval of ± 0.02, it is 92 participants.

CONCLUSIONS

Given that the majority of 16 studies included in this review reported ICCs of ≥ 0.90 and CVs of ≤ 5%, it can be concluded that the isometric mid-thigh pull maximum strength assessment has good-to-excellent test-retest reliability. The reliability is good-to-excellent for both absolute and relative peak force assessments and for both bilateral and unilateral isometric mid-thigh pull tests. The majority of studies did not find significant differences in peak force between testing sessions. From a practical perspective, we conclude that sport and exercise practitioners can use the bilateral and unilateral isometric mid-thigh pull tests as reliable measures of relative and absolute peak force. Additionally, practitioners should also consider that peak force values obtained in this test do not seem to change systematically between repeated measures.

Conflict of interest declaration

No potential conflict of interest was reported by the authors.

REFERENCES

1. Stone MH. Position statement: explosive exercises and training. Natl Strength Cond Assoc J. 1993; 15(3):7–15.
2. Narici MV, Hoppeler H, Kayser B, Landoni L, Claassen H, Gavardi C, Conti M, Cerretelli P. Human quadriceps cross-sectional area, torque and neural activation during 6 months strength training. Acta Physiol Scand. 1996; 157(2):175–186.
3. Stone MH, Bryant HO, Hornsby WG, Cunanan A, Mizuguchi S, Suarez DG, South M, Marsh DJ, Haff GG, Ramsey MW, Beckham GK, Santana HAP, Wagle JP, Stone ME, Pierce KC. Using the isometric mid-thigh pull in the monitoring of weightlifters: 25+ years of experience. UKSCA. 2019; 54:10–26.
4. Comfort P, Dos'Santos T, Beckham GK, Stone MH, Guppy SN, Haff GG. Standardization and methodological considerations for the isometric midthigh pull. Strength Cond J. 2019; 41(2):57–79.
5. Rodacki ALF, Boneti Moreira N, Pitta A, Wolf R, Melo Filho J, Rodacki CLN, Pereira G. Is Handgrip Strength a Useful Measure to Evaluate Lower Limb Strength and Functional Performance in Older Women? Clin Interv Aging. 2020; 15:1045–1056.
6. Suchomel TJ, Nimphius S, Stone MH. The Importance of Muscular Strength in Athletic Performance. Sports Med. 2016; 46(10):1419–1449.
7. Haff GG, Carlock JM, Hartman MJ, Kilgore JL, Kawamori N, Jackson JR, Morris RT, Sands WA, Stone MH. Force–time curve characteristics of dynamic and isometric muscle actions of elite women Olympic weightlifters. J Strength Cond Res. 2005; 19(4):741–748.
8. McGuigan MR, Winchester JB. The relationship between isometric and dynamic strength in collegiate football players. J Sports Sci Med. 2008; 7:101–105.
9. De Witt JK, English KL, Crowell JB, Kalogera KL, Guilliams ME, Nieschwitz BE, Hanson AM, Ploutz-Snyder LL. Isometric midthigh pull reliability and relationship to deadlift one repetition maximum. J Strength Cond Res. 2018; 32(2):528–533.
10. Thomas C, Comfort P, Chiang CY, Jones PA. Relationship between isometric mid-thigh pull variables and sprint and change of direction performance in collegiate athletes. J Trainology. 2015; 4:6–10.
11. Spiteri T, Nimphius S, Hart NH, Specos C, Sheppard JM, Newton RU. Contribution of strength characteristics to change of direction and agility performance in female basketball athletes. J Strength Cond Res. 2014; 28(9):2415–2423.
12. Townsend JR, Bender D, Vantrease WC, Hudy J, Huet K, Williamson C, Bechke E, Serafini PR, Mangine GT. Isometric midthigh pull performance is associated with athletic performance and sprinting kinetics in division I men and women's basketball players. J Strength Cond Res. 2019; 33(10):2665–2673.
13. Stone MH, Sands WA, Carlock J, Callan S, Dickie D, Daigle K, Cotton J, Smith SL, Hartman M. The importance of isometric maximum strength and peak rate-of-force development in sprint cycling. J Strength Cond Res. 2004; 18(4):878–884.
14. Kawamori N, Rossi SJ, Justice BD, Haff EE, Pistilli EE, O'Bryant HS, Stone MH, Haff GG. Peak force and rate of force development during isometric and dynamic mid-thigh clean pulls performed at various intensities. J Strength Cond Res. 2006; 20(3):483–491.
15. Carroll KM, Bernards JR, Bazyler CD, Taber CB, Stuart CA, DeWeese BH, Sato K, Stone MH. Divergent performance outcomes following resistance training using repetition maximums or relative intensity. Int J Sports Physiol Perform. 2019; 14(1):46–54.
16. Farney TM, MacLellan MJ, Hearon CM, Johannsen NM, Nelson AG. The effect of aspartate and sodium bicarbonate supplementation on muscle contractile properties among trained men. J Strength Cond Res. 2020; 34(3):763–770.
17. Marcus L, Soileau J, Judge LW, Bellar D. Evaluation of the effects of two doses of alpha glycerylphosphorylcholine on physical and psychomotor performance. J Int Soc Sports Nutr. 2017; 14:39.
18. Halperin I, Williams KJ, Martin DT, Chapman DW. The effects of attentional focusing instructions on force production

during the isometric midthigh pull. J Strength Cond Res. 2016; 30(4):919–923.
19. Thomas C, Jones PA, Comfort P. Reliability of the dynamic strength index in college athletes. Int J Sports Physiol Perform. 2015; 10(5):542–545.
20. Bartolomei S, Sadres E, Church DD, Arroyo E, Gordon JA III, Varanoske AN, Wang R, Beyer KS, Oliveira LP, Stout JR, Hoffman JR. Comparison of the recovery response from high-intensity and high-volume resistance exercise in trained men. Eur J Appl Physiol. 2017; 117(7):1287–1298.
21. Aben HGJ, Hills SP, Higgins D, Cooke CB, Davis D, Jones B, Russell M. The reliability of neuromuscular and perceptual measures used to profile recovery, and the time-course of such responses following academy rugby league match-play. Sports. 2020; 8(5):73.
22. Darrall-Jones JD, Jones B, Till K. Anthropometric and Physical Profiles of English Academy Rugby Union Players. J Strength Cond Res. 2015; 29(8):2086–2096.
23. Hornsby WG, Tice AL, Stone JD, Merrigan JJ, Hagen J, Wagle JP, Cunanan AJ, Stone MH. Changes in Maximal Strength and Home Run Performance in NCAA Division I Baseball Players Across 3 Competitive Seasons: A Descriptive Study. J Funct Morphol Kinesiol. 2021; 6(1):E4.
24. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998; 26(4):217–238.
25. Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. Sports Med. 2008; 38(4):297–316.
26. Comfort P, Jones PA, McMahon JJ, Newton R. Effect of knee and trunk angle on kinetic variables during the isometric midthigh pull: test-retest reliability. Int J Sports Physiol Perform. 2015; 10(1):58–63.
27. Comfort P, Dos'Santos T, Jones PA, McMahon JJ, Suchomel TJ, Bazyler C, Stone MH. Normalization of early isometric force production as a percentage of peak force during multijoint isometric assessment. Int J Sports Physiol Perform. 2020; 15(4): 478–482.
28. Dos'Santos T, Thomas C, Comfort P, McMahon JJ, Jones PA, Oakley NP, Young AL. Between-session reliability of isometric midthigh pull kinetics and maximal power clean performance in male youth soccer players. J Strength Cond Res. 2018; 32(12):3364–3372.
29. Dos'Santos T, Thomas C, Jones PA, McMahon JJ, Comfort P. The effect of hip joint angle on isometric midthigh pull kinetics. J Strength Cond Res 2017; 31(10):2748–2757.
30. Guppy SN, Brady CJ, Kotani Y, Stone MH, Medic N, Haff GG. The effect of altering body posture and barbell position on the between-session reliability of force-time curve characteristics in the isometric mid-thigh pull. Sports. 2018; 6(4):162.
31. Haines BR, Bourdon PC, Deakin G. Reliability of common neuromuscular performance tests in adolescent athletes. J Aust Strength Cond. 2016; 24(4):16–22.
32. James LP, Roberts LA, Haff GG, Kelly VG, Beckman EM. Validity and reliability of a portable isometric mid-thigh clean pull. J Strength Cond Res. 2017; 31(5):1378–1386.
33. Moeskops S, Oliver JL, Read PJ, Cronin JB, Myer GD, Haff GG, Lloyd RS. Within- and between-session reliability of the isometric midthigh pull in young female athletes. J Strength Cond Res. 2018; 32(7):1892–1901.
34. Sawczuk T, Jones B, Scantlebury S, Weakley J, Read D, Costello N, Darrall-Jones J, Stokes K, Till K. Between-day reliability and usefulness of a fitness testing battery in youth sport athletes: reference data for practitioners. Meas Phys Educ Exerc Sci. 2018; 22(1):11–18.
35. Sheppard JM, Chapman D, Taylor KL. An evaluation of a strength qualities assessment method for the lower body. J Aust Strength Cond. 2011; 19(2):4–10.
36. Suarez DG, Carroll KM, Slaton JA, Rochau KG, Davis MW, Stone MH. Utility of a shortened isometric midthigh pull protocol for assessing rapid force production in athletes. J Strength Cond Res. 2020. Doi: 10.1519/JSC.0000000000003774.
37. Thomas C, Dos'Santos T, Comfort P, Jones PA. Between-session reliability of common strength- and power-related measures in adolescent athletes. Sports 2017; 5(1):15.
38. Thomas C, Comfort P, Jones PA, Dos'Santos T. Between-session reliability of the unilateral stance isometric mid-thigh pull. J Aust Strength Cond. 2017; 25(5):6–10.
39. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009; 6(7):e1000097.
40. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability esearch. J Chiropr Med. 2016; 15(2):155–163.
41. Machin D, Campbell MJ, Walters SJ. Medical statistics. John Wiley & Sons Ltd, Chichester; 2007. pp. 203.
42. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010; 19(4):539–549.
43. Grgic J, Lazinica B, Schoenfeld BJ, Pedisic Z. Test-retest reliability of the one-repetition maximum (1RM) strength assessment: a systematic review. Sports Med Open. 2020; 6(1):31.
44. Beckham GK, Sato K, Santana HAP, Mizuguchi S, Haff GG, Stone MH. Effect of body position on force production during the isometric midthigh pull. J Strength Cond Res. 2018; 32(1):48–56.
45. Lehance C, Binet J, Bury T, Croisier JL. Muscular strength, functional performances and injury risk in professional and junior elite soccer players. Scand J Med Sci Sports. 2009; 19(2):243–251.
46. LaRoche DP, Cook SB, Mackala K. Strength asymmetry increases gait asymmetry and variability in older women. Med Sci Sports Exerc. 2012; 44(11):2172–2181.
47. Brady CJ, Harrison AJ, Comyns TM. A review of the reliability of biomechanical variables produced during the isometric mid-thigh pull and isometric squat and the reporting of normative data. Sports Biomech. 2020; 19(1):1–25.
48. Trevethan R. Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. Health Serv Outcomes Res Methodol. 2017; 17(2):127–143.
49. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med. 2002; 21(9):1331–1335.