©Journal of Sports Science and Medicine (2025) **24**, 543-554
http://www.jssm.org DOI: https://doi.org/10.52082/jssm.2025.543

Review article

Impact of Potential Moderating Factors on Absolute Test-Retest Reliability of Grip Strength Measurements in Healthy Populations: A Systematic Review with Meta-Analysis

Takashi Abe [1,2]✉, Jun Seob Song [3], Scott J. Dankel [4], Ricardo B. Viana [5], Akemi Abe [2] and Jeremy P. Loenneke [6]

[1] Graduate School of Health and Sports Science, Institute of Health and Sports Science & Medicine, Juntendo University, Chiba, Japan; [2] Division of Children's Health and Exercise Research, Institute of Trainology, Fukuoka, Japan; [3] Department of Counseling, Health and Kinesiology, Texas A&M University-San Antonio, San Antonio, TX, USA; [4] Department of Health and Exercise Science, Rowan University, Glassboro, NJ, USA; [5] Human Anatomy Laboratory, Institute of Physical Education and Sport, Federal University of Ceará, Fortaleza, Ceará, Brazil; [6] Department of Health, Exercise Science, and Recreation Management, Kevser Ermin Applied Physiology Laboratory, The University of Mississippi, Oxford, MS, USA

Abstract

Grip strength, a biomarker, can be measured at any age; however, its values vary daily for each individual, which impacts the assessment. Absolute test-retest reliability (i.e., minimal difference, MD) is commonly defined as the variation in absolute values of measurements taken by a single person or instrument on the same item under identical conditions. Nevertheless, the potential moderators of absolute repeatability in grip strength measurements have not yet been fully elucidated. We conducted a systematic review with meta-analysis to examine the influence of potential moderating factors on the absolute test-retest repeatability of grip strength measurements in healthy populations. PubMed, Scopus, and SPORTDiscus databases were searched up to January 2025 following the PRISMA guidelines, and 48 studies were included in this review. Age, test-retest interval, and device were used as potential moderating factors; however, sex and sports experience were excluded due to the limited number of published articles. We found considerable variation among studies reporting MD and percentage of MD to measured value (%MD) across each age group. The mean MD (%MD) values were 1.9 kg (25.4%) in young children (<7 years old), 2.5 kg (13.8%) in children (7-10 years old), 4.2 kg (17.1%) in adolescents (10 - 18 years old), 4.0 kg (11.6%) in young adults (18 - 35 years old), and 4.7 kg (16.7%) in older adults (>60 years old). Neither age [effect size [ES]: 0.015 (95% confidence interval [CI]: -0.004, 0.035; $p = 0.113$) for MD and ES: -0.025 (95% CI: -0.089, 0.039; $p = 0.439$) for %MD], test-retest interval [ES: 0.006 (95% CI: -0.002, 0.013; $p = 0.143$) for MD and ES: 0.022 (95% CI: -0.001, 0.046; $p = 0.065$) for %MD] nor handgrip device ($p = 0.752$ for MD and $p = 0.334$ for %MD) served as significant moderators of MD and %MD reliability. Due to the limited number of studies, sex and sports experience were excluded from the analysis; as a result, their impacts remain unknown.

Key words: Dynamometer, handgrip, reproducibility, peak muscular strength.

Introduction

Grip (or handgrip) strength is an extensively used biomarker in research and clinical practice within the health, sports science, nutrition, and medical fields (Abe et al., 2022; Bohannon, 2015; Bohannon, 2019; Norman et al., 2011). An online literature search (i.e., grip strength as a keyword) using PubMed identified over 48,000 publications at a rate of over 3,500 per year in the last five years. These publications include scientific literature and guidelines discussing grip strength's reference values in each age group in both sexes (Abe et al., 2016; Hanten et al., 1999; Ramirez-Velez et al., 2021) and its association with current and future health (Celis-Morales et al., 2017; Peralta et al., 2023; Rantanen et al., 2003). For example, grip strength increases dramatically from preschool children to young adults, maintains stability in middle age, and then declines in old age (Abe et al., 2024; Loenneke et al., 2024; Stenholm et al., 2012). In children and adolescents, grip strength may be a valuable indicator of bone health that improves with growth (Saint-Maurice et al., 2018). Grip strength is also used as a criterion for diagnosing sarcopenia in middle-aged and older men and women (Cruz-Jentoft et al., 2019). However, individual grip strength varies daily, and the degree of these changes may differ depending on age, sex, device type, and sports experience (i.e., athletes). For example, assuming similar daily variation in each individual, the absolute test-retest reliability of grip strength measurements is expected to differ, being lower in children and older adults with low grip strength levels than in younger adults. The same is true for both men and women in adolescence and beyond. Additionally, athletes in sports may be better equipped to consistently exert maximum muscle strength. Thus, measurement error should be considered when comparing the measured grip strength values with the evaluation or diagnosis criteria.

As is the case with many studies assessing reliability in the exercise science literature, the intraclass correlation coefficient (ICC) appears to be the preferred method for reporting the reliability of grip strength measurements (Bobos et al., 2020; Bohannon, 2017). One major limitation with reporting ICC values is that they are entirely dependent on the heterogeneity of the sample included in the reliability assessment (i.e. between subject variability) given that the ICC is calculated with the following formula:

Received: 18 April 2025 / Accepted: 25 June 2025 / Published (online): 01 September 2025

$$ICC = \frac{\text{betweeen subject variability}}{\text{between subject variability} + \text{within subject variability}}$$

Thus, if the sample recruited is very homogenous (low between subject variability), the ICC values will likely be small demonstrating poor reliability, even if the absolute test-retest reliability is good. On the contrary, if the sample recruited is very heterogenous (high between subject variability), the ICC values will likely be high demonstrating good reliability, even if the absolute test-retest reliability is not good (Weir, 2005). While there are certainly instances where relative reliability may be important (i.e. epidemiologic studies, correlations between variables, etc.), often absolute reliability (i.e., standard error of measurement (SEM) and minimal difference (MD)) is preferred and more useful (Weir, 2005). For example, whenever repeated measures are used (i.e. training interventions), individuals are only compared to themselves, so factoring in between subject variability is not important. However, good absolute test-retest reliability will reduce the variability (i.e. error) amongst repeated measures, improving the ability to detect true changes by reducing the denominator in the test statistic. The same holds true for detecting differences between groups, where better absolute test-retest reliability within each group is preferred to lower the pooled standard error. Despite the importance of establishing absolute test-retest reliability of grip strength, it is unclear to what extent it varies with age and other physical factors and measurement methods (e.g., time interval, device used). Therefore, this systematic review with meta-analysis examined the impact of potential moderators on absolute test-retest repeatability of grip strength measurements in healthy populations.

Methods

We conducted this systematic review according to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement (Page et al., 2021). The study was pre-registered (January 8, 2025) in the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42025635760).

English-language searches of the electronic databases PubMed, Scopus, and SPORTDiscus were conducted from inception to January 9, 2025, by a researcher (J.S.S.). Relevant articles were retrieved from electronic databases combining the following terms: (handgrip OR grip) AND (reliability OR retest OR reproducibility OR repeatability). Initially, all files were extracted from databases in either RIS (Scopus and SPORTDiscus) or NBIB (PubMed) format. The files were then uploaded into Rayyan software, where duplicates were eliminated. Subsequently, two reviewers (T.A. and J.S.S.) independently checked the titles and abstracts of identified articles for relevance. The reviewers then independently reviewed the full text of potentially eligible papers. Any disagreements between the reviewers on inclusion were resolved by a consensus between both reviewers. Additional articles were identified via hand-searching and reviewing the reference list of relevant papers. The study selection process is summarized using the PRISMA flow diagram (Figure 1).

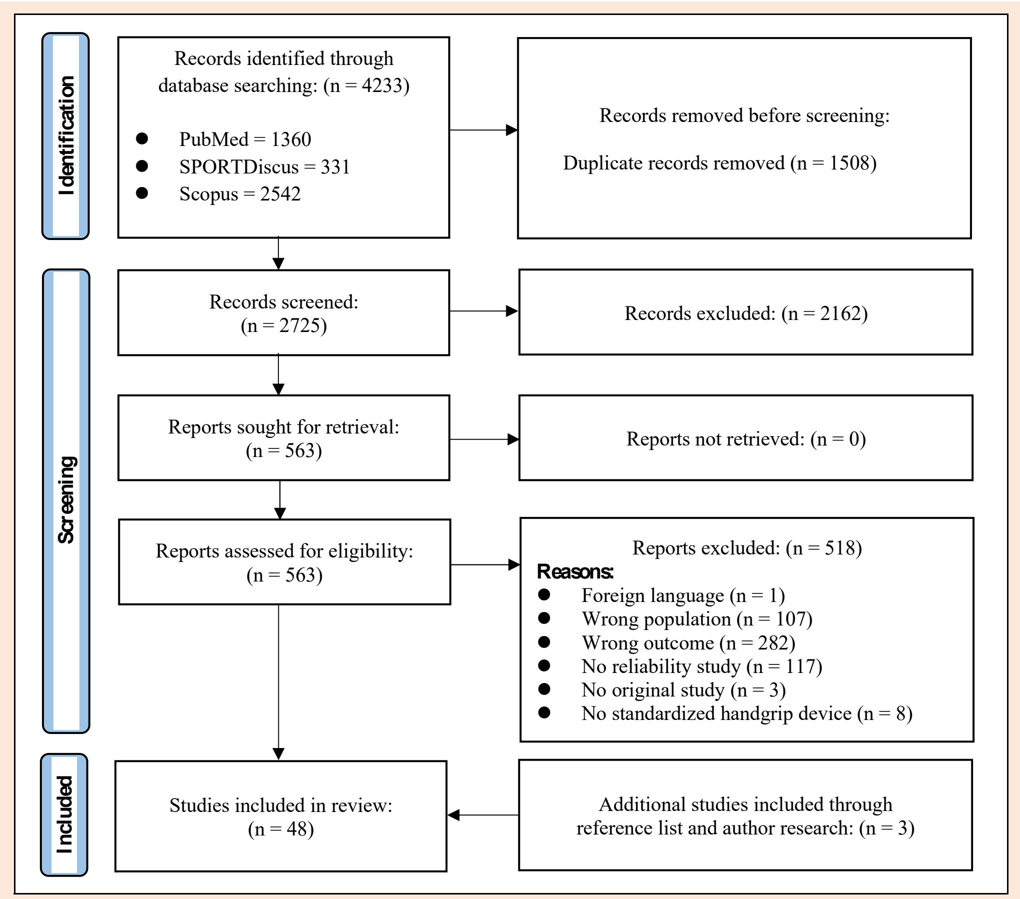

Figure 1. PRISMA flowchart of outcomes of the search strategy.

To be included in this systematic review, studies were required to fulfill the following criteria: (1) a published original study written in English language; (2) healthy participants with no restrictions on age, sex, and physical activity/training status; (3) measured maximum handgrip strength using a standardized handgrip dynamometer at both test and retest, with the same investigator conducting the measurements (i.e., intra-observer); and (4) reported absolute reliability (i.e., MD) or provided data needed to calculate absolute reliability (e.g., standard error of measurement, standard deviation of test-retest mean difference). Studies were included if they targeted healthy participants based on the title and abstract of the articles, but were excluded if they targeted participants with any diseases. When reliability information was not available in the title and abstract, we examined the characteristics of the participants in the articles. If there was no mention of a study participant's chronic diseases, they were considered "healthy individuals." If a study did not report absolute reliability, then we calculated the MD using: the standard deviation of test-retest mean difference (SDd) (*Equation 1*); and SEM (*Equation 2*).

$$\text{Equation 1: } MD = SDd \times 1.96$$
$$\text{Equation 2: } MD = SEM \times 1.96 \times \sqrt{2}$$

If a study reported the test-retest pooled SD with intraclass correlation (ICC), we first calculated the SEM using *Equation 3*. The calculated SEM was then used to determine the MD using *Equation 2*. For studies that did not report the test-retest pooled SD, we used the SD of test (pre-test), assuming that test-retest pooled SD and test SD would be similar, as both tests were completed by the same individuals.

$$\text{Equation 3: } SEM = test\ \&\ retest\ pooled\ SD\ (or\ test\ SD) \times \sqrt{(1 - ICC)}$$

The percentage MD (%MD) was calculated following *Equation 4*. When the study did not report test-retest pooled values, the mean grip strength was calculated as average using the test and retest values.

$$\text{Equation 4: } \%MD = MD \div \text{mean test \& retest grip strength} \times 100$$

The following study characteristics were extracted: authors, publication year, participants' characteristics (age, sex, and health status), sample size, handgrip device, time interval between test and retest, handgrip strength at test and retest (with SD), test and retest mean difference (with SD), ICC, SEM, and MD. Two researchers (T.A. and J.S.S.) extracted these data manually, with disagreement resolved by consensus between both researchers. To standardize grip strength values to kilogram units, those reported in newtons were converted as follows: 1 N = 0.10197 kg. In studies reporting the sum of grip strength values for both the left and right hands, the MD value was divided by 2 for statistical processing, as many studies used grip strength values and MD values for one hand.

A modified version of the critical appraisal tool was utilized to evaluate the methodological quality of the studies included in this review (Brink and Louw, 2011), and two researchers (S.J.D. and J.S.S.) independently evaluated the included studies. Seven relevant items were extracted from the modified checklist: 1) Subject characteristics were clearly described, 2) The competence of the raters was explicitly detailed, 3) Raters were blinded to their previous findings, 4) The time interval between repeated measures was appropriate, 5) The execution of the test was described in sufficient detail to allow replication, 6) Study participant's withdrawals were clearly explained, 7) The statistical methods were suitable for the study's objectives. Other items were not included as they were not considered relevant for this review. The score for each item was determined as follows: 1 = yes; 0 = no. Consequently, the maximal possible score was 7.

To account for the dependency of multiple effect sizes nested within individual studies, a multi-level model was employed using the metafor package (version 4.6 - 0) in RStudio (version: 2024.12.1 + 563) (Assink and Wibbelink, 2016). Three models were computed to assess 1) the MD, 2) the %MD, and 3) systematic bias (test 2 - test 1). Since the effect sizes of interest were computed off variability statistics to assess reliability, each of the studies was weighted based on the sample size as we have done previously (Dankel et al., 2019). Three moderating variables were also assessed for the MD and %MD which included: 1) age (continuous: years), 2) time interval between test-retest (continuous: days), and handgrip device used (categorical: Jamar, Takei, other) to determine their influence on reliability. For systematic bias (calculated as the change from test 1 to test 2), age was used as a moderator to determine if children or younger adults may have experienced a greater learning effect. Sex and sports experience were not used as moderating variables for any of the analyses, given the limited number of studies that assessed males and females separately and the limited number of studies assessing athletes. In summarizing the results (Tables 1, 2, and 3), the following age ranges were used for age categorization: young children (under 7 years old), children (between 7 - 10 years old), adolescents (between 10 - 18 years old), young adults (between 18 - 35 years old), middle-aged adults (between 36 - 60 years old), and older adults (>60 years old). Statistical significance was set at $p < 0.05$.

Results

Included studies and participant and research protocol characteristics

The original article search yielded 4,233 studies. Three additional studies were identified from the reference lists of the included articles. After removing duplicates and eliminating articles based on the eligibility criteria, 48 studies (Abe et al., 2018; Abe et al., 2019; Abe et al., 2022; Amado-Pacheco et al., 2019; Anstey et al., 1997; Balogun et al., 1991; Beauchamp et al., 2021; Biasini et al., 2023; Bohannon, 2006; Bohannon and Schaubert, 2005; Bohannon et al., 2011; Boshnjaku et al., 2021; Cadenas-Sanchez et al., 2016; Cildan Uysal et al., 2022; Dugdale et al., 2019; Espana-Romero et al., 2010; Essendrop et al., 2001; Fernandez-Santos et al., 2016; Ferreira et al., 2021; Gasior et al., 2020; Gerodimos, 2012; Gerodimos and Karatrantou,

Table 1. Study characteristics and absolute reliability in children and adolescents.

Author (Year)	Age (yr)	Sex	Device	Interval	Arm	HGS test1	HGS test2	MD (kg)	% MD
Young Children (< 7 years old)									
Svensson et al (2008)	6 years	19BG	Grippit	7 days	NR	7.65 (2.24)	7.85 (2.65)	1.36	17.5
Sanchez-Delgado et al (2015)	(3-5)	32B	Takei	3 hours	NR	6.6 (3.0)	6.2 (2.6)	2.52	39.4
	(3-5)	24G				6.6 (3.1)	6.7 (3.0)	2.17	32.7
Cadenas-Sanchez et al (2016)	4.90 (0.86)	92B	Takei	14 days	Avg of both hands	8.40 (2.40)	8.02 (2.56)	2.59	31.5
	4.82 (0.79)	69G				7.24 (2.34)	7.22 (2.21)	2.37	32.8
Amado-Pacheco et al (2019)	4.04 (0.82)	48B	NR	14 days	Avg of both hands	8.19 (2.46)	8.10 (2.39)	1.14	14.0
	3.95 (0.82)	42G				7.30 (1.95)	7.35 (1.92)	1.41	19.3
Abe et al (2022)	6.1 (0.3)	8B 5G	Takei	7 days	Right	10.3 (1.4)	10.1 (1.1)	1.60	15.8
King-Dowling et al (2024)	4.7 (0.6)	22B 20G	Takei	2-3 weeks	Dominant	7.4 (1.6)	7.8 (1.6)	2.16	28.4
					Nondominant	7.7 (1.5)	7.8 (1.5)	1.76	22.8
Children (7-10 years)									
Espana-Romero et al (2010)	(6-11)	58BG	Takei	7 days	Sum of both hands	29.9 (4.9)	28.4 (3.8)	4.31	14.8
Gerodimos (2012)	9.85 (0.70)	30B [BB]	Jamar	24 hours	Preferred	20.06 (4.67)	20.32 (4.81)	3.23	16.0
					Non-preferred	19.78 (4.59)	19.92 (4.59)	2.37	12.0
Gerodimos & Karatrantou (2013)	9.49 (0.96)	27B [WR]	Jamar	24 hours	Preferred	22.33 (3.37)	22.11 (3.66)	2.55	11.5
					Non-preferred	21.78 (3.30)	21.59 (3.42)	2.74	12.7
Fernandez-Santos et al (2016)	8.7 (1.8)	98B 82G	Takei	7 days	Sum of both hands	31.5 (10.1)	31.3 (9.9)	5.25	16.7
Gasior et al (2020)	7~9	69B 68G	Jamar	24 hours	Dominant	13.25 (3.28)	13.75 (3.34)	1.78	13.2
Adolescents (10-18 years old)									
Svensson et al (2008)	10 years	20BG	Grippit	7 days	NR	15.7 (3.98)	17.0 (4.89)	5.68	34.8
	14 years	19BG				32.3 (8.66)	33.1 (8.57)	4.74	14.5
Ortega et al (2008)	13.7 (0.8)	69B	Takei	2 weeks	Avg of both hands	31.2 (6.4)	31.5 (6.9)	4.90	15.6
	13.6 (0.8)	54G				26.1 (5.1)	26.1 (4.9)	3.53	13.5
Espana-Romero et al (2010)	(12-18)	80BG	Takei	7 days	Sum of both hands	50.5 (14.6)	51.9 (15.3)	9.60	18.8
Gerodimos (2012)	14.37 (0.61)	30B [BB]	Jamar	24 hours	Preferred	42.10 (9.44)	42.67 (9.15)	3.14	7.4
					Non-preferred	41.27 (8.41)	41.81 (8.73)	3.25	7.8
Gerodimos & Karatrantou (2013)	14.60 (0.50)	27B [WR]	Jamar	24 hours	Preferred	47.07 (8.32)	47.26 (8.56)	2.94	6.2
					Non-preferred	46.67 (9.2)	46.63 (9.23)	3.60	7.7
Ramirez-Velez et al (2015)	12.8 (2.4)	124B	Takei	7 days	Avg of both hands	19.6 (8.9)	19.0 (8.5)	3.92	20.3
	12.8 (2.5)	105G				16.9 (5.1)	16.5 (5.4)	3.14	18.8
Dugdale et al (2019)	NR for each group, overall 13.5 (1.8) [SO]	26B	Takei	7-14 days	Dominant	17.8 (2.6)	17.7 (2.9)	1.91	10.7
		51B				18.1 (3.6)	18.7 (3.3)	4.11	22.4
		75B				21.5 (4.5)	22.1 (4.8)	4.83	22.2
		59B				25.3 (5.3)	25.9 (5.8)	4.87	19.0
		81B				33.2 (7.5)	33.4 (7.2)	5.88	17.7
		46B				37.7 (7.2)	38.3 (6.3)	5.28	13.9
		35B				37.5 (7.3)	37.8 (7.0)	7.01	18.6
Gasior et al (2020)	10~13	87B 82G	Jamar	24 hours	Dominant	22.68 (5.60)	22.70 (5.50)	2.41	10.6
O'Keeffe et al (2020)	13.44 (0.35)	19B, 26G	Takei	7 days	Avg of both hands	21.5 (4.1)	21.5 (4.5)	2.16	10.0
	13.42 (0.32)	20B, 21G				25.3 (5.4)	25.0 (5.3)	1.76	7.0
Trajkovic et al (2024)	12.2 (0.4)	32B 24G	Jamar	5 days	Right	20.50 (4.89)	19.32 (4.91)	5.25	26.4
					Left	18.64 (4.07)	18.01 (4.39)	5.17	28.2
			Takei		Right	20.81 (4.84)	20.16 (5.34)	5.53	27.0
					Left	19.22 (4.26)	19.01 (4.94)	5.54	29.0

B, boys; G, girls; Avg, average; NR, not reported; MD, minimal difference; %MD, percentage of a minimal difference to the measured value; HGS, handgrip strength (unit in kilograms); BB, basketball players; WR, wrestlers; SO, soccer players

2013; Gil et al., 2022; Hamilton et al., 1992; Jenkins and Cramer, 2017; Karatrantou et al., 2020; Kieser et al., 2025; King-Dowling et al., 2024; Legg et al., 2020; Lemmink et al., 2001; Leszczak et al., 2024; Maurya et al., 2023; O'Keeffe et al., 2020; Ortega et al., 2008; Petersen et al., 2015; Plant et al., 2016; Ramirez-Velez et al., 2015; Sanchez-Delgado et al., 2015; Savva et al, 2013; Suzuki et al., 2019; Svensson et al., 2008; Tan et al., 2001; Trajkovic et al., 2024; Tsang, 2005; Venegas-Carro et al., 2022; Villafane et al., 2016; Walamies and Turjanmaa, 1993; Ward and Adams, 2007) were included in this review. Of those studies, 16 included 42 data points assigned to children and adolescents (852 boys, 294 girls, and 879 mixed) (Abe et al., 2022; Amado-Pacheco et al., 2019; Cadenas-Sanchez et al., 2016; Dugdale et al., 2019; Espana-Romero et al., 2010; Fernandez-Santos et al., 2016; Gasior et al., 2020; Gerodimos, 2012; Gerodimos and Karatrantou, 2013; King-Dowling et al., 2024; O'Keeffe et al., 2020; Ortega et al., 2008; Ramirez-Velez et al., 2015; Sanchez-Delgado et al., 2015; Svensson et al., 2008; Trajkovic et al., 2024)

(Table 1), 25 included 50 data points assigned to young and middle-aged adults (98 men, 109 women, and 1,244 mixed) (Abe et al., 2018; Abe et al. 2019; Balogun et al., 1991; Beauchamp et al., 2021; Biasini et al., 2023; Bohannon, 2006; Bohannon et al., 2011; Boshnjaku et al., 2021; Cildan Uysal et al., 2022; Essendrop et al., 2001; Gerodimos, 2012; Gil et al., 2022; Hamilton et al., 1992; Karatrantou et al., 2020; Kieser et al., 2025; Leszczak et al., 2024; Maurya et al., 2023; Petersen et al., 2015; Plant et al., 2016; Savva et al., 2013; Tan et al., 2001; Tsang, 2005; Venegas-Carro et al., 2022; Walamies and Turjanmaa, 1993; Ward and Adams, 2007) (Table 3), and 12 included 23 data points assigned to older adults (166 men, 292 women, and 1,046 mixed) (Abe et al., 2018; Anstey et al., 1997; Beauchamp et al., 2021; Bohannon and Schaubert, 2005; Boshnjaku et al., 2021; Ferreira et al., 2021; Gil et al., 2022; Jenkins and Cramer, 2017; Legg et al., 2020; Lemmink et al., 2001; Suzuki et al., 2019; Villafane et al., 2016) (Table 2). The following studies (Bohannon, 2006; Bohannon et al., 2011; Kieser et al., 2025; Plant et al., 2016; Tsang, 2005) included participants spanning a broad range of adults (young, middle-aged, and older) and were therefore included in Table 3.

The main dynamometers used to measure grip strength were Jamar (Abe et al., 2019; Bohannon and Schaubert, 2005; Bohannon et al., 2011; Boshnjaku et al., 2021; Essendrop et al., 2001; Gasior et al., 2020; Gerodimos, 2012; Gerodimos and Karatrantou,2013; Hamilton et al., 1992; Jenkins and Cramer, 2017; Karatrantou et al., 2020; Legg et al., 2020; Lemmink et al., 2001; Savva et al., 2013; Trajkovic et al., 2024; Tsang, 2005; Venegas-Carro et al., 2022; Villafane et al., 2016; Ward and Adams, 2007) and Takei (Abe et al., 2018; Abe et al., 2019; Abe et al., 2022; Cadenas-Sanchez et al., 2016; Dugdale et al., 2019; Fernandez-Santos et al., 2016; Gerodimos and Karatrantou, 2013; King-Dowling et al., 2024; O'Keeffe et al., 2020; Ortega et al., 2008; Petersen et al., 2015; Ramirez-Velez et al., 2015; Sanchez-Delgado et al., 2015; Suzuki et al., 2019; Tan et al., 2001; Trajkovic et al., 2024). Other studies used different types of dynamometers, such as Grippit (Svensson et al., 2008), JTECH (Biasini et al., 2023; Plant et al., 2016), and MicroFET (O'Keeffe et al., 2020). Two studies did not report the type of dynamometers (Amado-Pacheco et al., 2019; Beauchamp et al., 2021).

The most commonly used test-retest intervals were 24 hours (Abe et al., 2018; Abe et al., 2019; Bohannon, 2006; Gasior et al., 2020; Gerodimos, 2012; Gerodimos et al., 2013; Tan et al., 2001; Ward and Adams, 2007) or 7 days (Abe et al., 2022; Beauchamp et al., 2021; Espana-Romero et al., 2010; Essendrop et al., 2001; Fernandez-Santos et al., 2016; Ferreira et al., 2021; Gil et al., 2022; Hamilton et al., 1992; Lemmink et al., 2001; O'Keeffe et al., 2020; Petersen et al., 2015; Ramirez-Velez et al., 2015; Savva et al., 2013; Svensson et al., 2008; Venegas-Carro et al., 2022; Villafane et al., 2016), with several studies using 2 weeks (Amado-Pacheco et al., 2019; Cadenas-Sanchez et al., 2016; Hamilton et al., 1992; Leszczak et al., 2024; Ortega et al., 2008). In 41 of the 48 studies, the test-retest interval was 2 weeks or less. Nine studies had a range of test-retest intervals that were not consistent, such as within 7 days (Balogun et al., 1991; Biasini et al., 2023; Maurya et al., 2023) or 2 - 10 days (Kieser et al., 2025).

Table 2. Study characteristics and absolute reliability in older adults (>60 years old).

Author (Year)	Age (yr)	Sex	Device	Interval	Arm	HGS test1	HGS test2	MD (kg)	% MD
Anstey et al. (1997)	67.92 (4.89)	50W	Spedly manual	3 months	Right	25.52 (4.55)	26.15 (5.91)	5.50	21.3
					Left	22.63 (4.16)	23.48 (5.03)	4.61	20.0
Lemmink et al. (2001)	65.8 (7.03)	68M	Jamar	7 days	Preferred	43.9 (8.00)	44.7 (8.45)	7.53	17.0
	66.1 (6.75)	83W				28.1 (5.30)	29.3 (5.55)	6.17	21.5
Bohannon & Schaubert (2005)	75.0 (5.9)	4M 17W	Jamar	12 weeks	Right	26.70 (7.39)	25.83 (7.05)	5.94	22.6
					Left	24.71 (7.69)	24.36 (7.44)	4.51	18.4
Villafane et al. (2016)	67.5 (10.2)	6M 9W	Jamar	7 days	Dominant	25.8 (9.9)	25.8 (9.9)	0.67	2.6
					Nondominant	24.4 (10.5)	24.3 (10.6)	2.22	9.1
Jenkins & Cramer (2017)	76.8 (6.3)	98M	Jamar	12 weeks	NR	29.76 (9.27)	29.62 (8.83)	5.49	18.5
				24 weeks		30.42 (9.09)	30.61 (8.98)	5.07	16.6
	75.9 (6.6)	159W		12 weeks		17.60 (6.10)	17.63 (6.08)	3.38	19.2
				24 weeks		18.17 (5.82)	18.27 (5.72)	2.66	14.6
Abe et al. (2018)	72 (3.8)	34M 46W	Takei	1 year	Right	31.8 (7.9)	29.5 (6.8)	6.4	21.1
Suzuki et al. (2019)	≥65	197M 21W	Takei	NA	Dominant	26.3 (6.8)	26.0 (6.8)	3.94	15.1
Legg et al. (2020)	71 (10)	6M 11W	Jamar	48 hours	Dominant	34.7 (15.1)	35.1 (14.4)	5.92	17.0
					Nondominant	32.6 (16.6)	33.6 (16.3)	4.60	13.9
Beauchamp et al. (2021)	69 (3)	29M 521W	NR	7 days	Dominant	37.08 (12.79)	39.08 (13.23)	10.64	28.0
	81 (4)	29M 20W				29.41 (9.28)	36.65 (12.63)	5.16	15.6
Boshnjaku et al. (2021)	70.7 (6.1)	22M 39W	Jamar	7-97 days	Dominant	29.2 (9.2)	28.8 (9.3)	5.29	18.3
Ferreira et al. (2021)	84.5 (6.5)	15M 28W	Baseline Smedley	7 days	Dominant	18.9 (5.9)	18.9 (5.8)	3.49	18.5
					Non-dominant	17.1 (5.6)	17.4 (5.4)	2.91	16.9
Gil et al. (2022)	70.5 (5.0)	6M 6W	Straingauge	7 days	Right	29.5 (8.4)	29.2 (8.0)	2.27	7.7
					Left	26.9 (7.8)	28.5 (7.1)	2.92	10.5

M, men; W, women; Avg, average; NR, not reported; MD, minimal difference; %MD, percentage of a minimal difference to the measured value; HGS, handgrip strength (unit in kilograms)

Table 3. Study characteristics and absolute reliability in young and middle-aged adults.

Author (Year)	Age (yr)	Sex	Device	Interval	Arm	HGS test1	HGS test2	MD (kg)	% MD
Young Adults (18-35 years old)									
Balogun et al. (1991)	23.7 (2.4)	60M	Harpenden	within 7 days	Right	36.7 (7.7)	36.2 (7.3)	4.23	11.6
					Left	35.4 (6.7)	35.0 (6.8)	4.59	13.0
		60W			Right	24.1 (5.2)	24.6 (5.0)	3.76	15.4
					Left	22.1 (4.5)	22.1 (4.2)	3.46	15.7
Hamilton et al. (1992)	23.8 (4.9)	29W	Jamar	7 days	Right	63.09 (9.43)*	69.21 (8.08)*	9.19	13.9
				14 days		63.09 (9.43)*	66.23 (10.50)*	11.55	17.9
Gerodimos (2012)	26.06 (5.57)	30M [BB]	Jamar	24 hours	Preferred	66.49 (9.33)	66.71 (9.68)	4.57	6.9
					Non-preferred	65.68 (9.16)	65.95 (9.26)	4.06	6.2
Savva et al. (2013)	21~26	10M 9W	Jamar	7 days	Dominant	35.6 (12.1)	36.2 (12.7)	6.45	18.0
Petersen et al. (2015)	25 (2)	8M	Takei	7 days	Right	50.8 (10.6)	(pooled)	5.43	10.7
					Left	48.0 (10.7)	(pooled)	6.93	14.4
Karatrantou et al. (2020)	18.5 (3.4)	14M 6W	Jamar	3 days	Preferred	40.45 (8.70)	40.65 (8.80)	2.66	6.6
					Non-preferred	39.85 (9.42)	40.00 (8.81)	4.80	12.0
Boshnjaku et al. (2021)	22.6 (3.7)	31M 26W	Jamar	7-97 days	Dominant	41.5 (10.7)	42.3 (11.0)	5.29	12.6
Venegas-Carro et al. (2022)	24.2 (2.2)	10M 7W	Jamar	7 days	Preferred	47.7 (12)	47.9 (12)	3.1	6.6
				9 weeks		47.7 (12)	48.4 (14)	5.7	11.8
Gil et al. (2022)	22.7 (2.8)	6M 6W	Straingauge	7 days	Right	35.0 (9.8)	35.8 (10.1)	4.78	13.5
					Left	33.0 (8.9)	34.0 (9.0)	3.51	10.5
Cildan Uysal et al. (2022)	22.32 (0.79)	19M 31W	K-Force Grip	48 hours	Right (standing)	25.92 (8.49)	25.86 (8.38)	2.68	10.4
					Left (standing)	23.98 (8.01)	23.75 (7.3)	3.22	13.5
					Right (sitting)	25.67 (8.36)	25.58 (8.25)	2.84	11.1
					Left (sitting)	23.69 (7.43)	23.64 (7.45)	3.32	14.0
Biasini et al. (2023)	24.4 (1.4)	10M 10W	JTECH	2-5 days	Dominant	45.4 (13.7)	44.4 (13.7)	6.90	15.4
Maurya et al. (2023)	21 (3)	20W	Dynamo Torque Analyzer	2-7 days	Dominant	18.8 (4.5)	18.2 (4.6)	3.37	18.2
Leszczak et al. (2024)	22.2 (1.46)	44M 78W	Biometrics E-Link EP9	2 weeks	Right (rater #1)	35.36 (11.69)	35.44 (11.77)	2.16	6.1
					Left (rater #1)	30.47 (9.89)	30.50 (9.81)	1.65	5.4
					Right (rater #2)	35.58 (11.66)	35.72 (11.70)	2.08	5.8
					Left (rater #2)	30.06 (9.87)	30.01 (9.68)	1.98	6.6
Young and Middle-aged Adults									
Walamies & Turjanmaa (1993)	23-49	13M 27W	Straingauge	1-2 months	Dominant	36.9 (11.2)	37.5 (11.2)	4.39	11.8
Essendrop et al. (2001)	35 (6.9)	6M 13W	Jamar	7 days	Right	39.8 (10.2)	41.5 (10.5)	4.00	9.8
Tan et al. (2001)	34.3 (8.2)	12M, 9W [BL]	Takei	24 hours	Bowling hand	38.2 (8.8)	37.4 (7.9)	6.45	17.1
Ward & Adams (2007)	19~47	9M 21W	Jamar	24 hours	Dominant	32.77 (10.43)	32.2 (9.85)	6.70	20.6
					Non-dominant	30.79 (9.69)	29.94 (8.47)	6.53	21.5
			MIE		Dominant	30.02 (8.17)	28.47 (8.91)	6.88	23.5
					Non-dominant	27.56 (8.13)	27.55 (8.23)	5.76	20.9
			Lafayette		Dominant	29.41 (12.04)	28.11 (12.38)	9.37	32.6
					Non-dominant	27.84 (10.97)	26.11 (10.69)	7.61	28.2
Abe et al. (2018)	54 (6)	9M 9W	Takei	24 hours	Right	40.1 (10.9)	38.7 (11.4)	3.97	10.1
Abe et al. (2019)	46.4 (13.2)	10M 10W	Takei	24 hours	Right	39.5 (10.8)	38.5 (10.8)	3.5	8.9
			Jamar			42.6 (12.8)	41.5 (12.8)	4.6	10.9
Beauchamp et al. (2021)	58 (4)	18M 30W	NR	7 days	Dominant	36.65 (12.63)	37.56 (12.84)	6.07	16.4
Young, Middle-age and Older Adults									
Tsang et al. (2005)	37.8 (10.9)	226M 322W	Jamar	3 days	Dominant	36.2 (10.4)	35.7 (10.3)	6.10	17.0
					Nondominant	33.7 (10.0)	33.3 (10.0)	6.10	18.2
Bohannon (2006)	38.0 (15.6)	14M, 16W	MicroFET	24 hours	Right	44.1 (15.3)	44.3 (16.1)	5.85	13.2
					Left	42.2 (14.1)	41.8 (14.7)	4.79	11.4
Bohannon et al. (2011)	45.7 (23.5)	14M, 14W	Jamar	4-10 days	Right	38.2 (10.0)	38.8 (10.8)	6.94	15.3
					Left	35.9 (9.8)	36.5 (10.2)	5.88	19.2
Plant et al. (2016)	(23-67)	10M 15W	JTECH (Manual)	11.3 (10.6) weeks	NR	27.5 (9.6)	26.7 (8.9)	5.32	19.6
			JTECH (Electronic)			28.4 (10.2)	27.9 (9.4)	5.66	20.1
Kieser et al. (2025)	48.0 (20.2)	39M 61W	Biopac	2-10 days	Max value of both hands	26.9 (9.7)	27.1 (9.6)	4.63	17.1

B, boys; G, girls; M, men; W, women; Avg, average; NR, not reported; MD, minimal difference; %MD, percentage of a minimal difference to the measured value; HGS, handgrip strength (unit in kilograms); BB, basketball players; BL, 10-pin bowlers.

*The unit of measured grip strength was pounds per square inch. Thus, this study was not included in the minimal differences (MD) analysis.

Table 4. Absolute reliability (minimal difference and percentage of minimal difference to the measured value) of grip strength measurements according to age, sex, and hand dominance.

		# of Studies	Minimal Difference Mean [Min, Max]	% Minimal Difference Mean [Min, Max]
Age Group	Young Children	6	1.9 [1.14, 2.59]	25.4 [14.0, 39.4]
	Children	5	2.5 [1.78, 3.23]	13.8 [11.5, 16.7]
	Adolescents	10	4.2 [1.76, 7.01]	17.1 [6.2, 34.8]
	Young Adults	13	4.0 [1.65, 6.93]	11.6 [5.4, 18.2]
	Young & Middle-aged Adults	7	5.8 [3.97, 9.37]	17.9 [8.9, 32.6]
	Older Adults	12	4.7 [0.67, 10.64]	16.7 [2.6, 28.0]
Sex	Boys	5	3.0 [1.14, 4.90]	24.2 [14.0, 39.4]
	Girls	5	2.5 [1.41, 3.53]	23.4 [13.5, 32.8]
	Men	3	5.4 [4.23, 7.53]	15.3 [11.6, 18.5]
	Women	3	3.9 [2.66, 6.17]	17.3 [14.6, 21.5]
Hand	Dominant	11	3.4 [0.67, 6.1]	12.5 [6.6, 28.4]
	Non-dominant	11	3.5 [1.76, 6.1]	12.7 [6.2, 22.8]
	Right	11	4.4 [2.08, 6.97]	14.5 [5.8, 27.0]
	Left	11	4.1 [1.65, 6.93]	15.3 [5.4, 29.0]

Min, minimum; Max, maximum

Assessment of methodological quality

The mean score was 4.2 out of 7 (range: 2 - 7), indicating a methodological quality rating that varied from low to high (Supplementary Table 1). Seventeen of the 48 studies scored 5 or higher, while 10 received scores of 2 or 3.

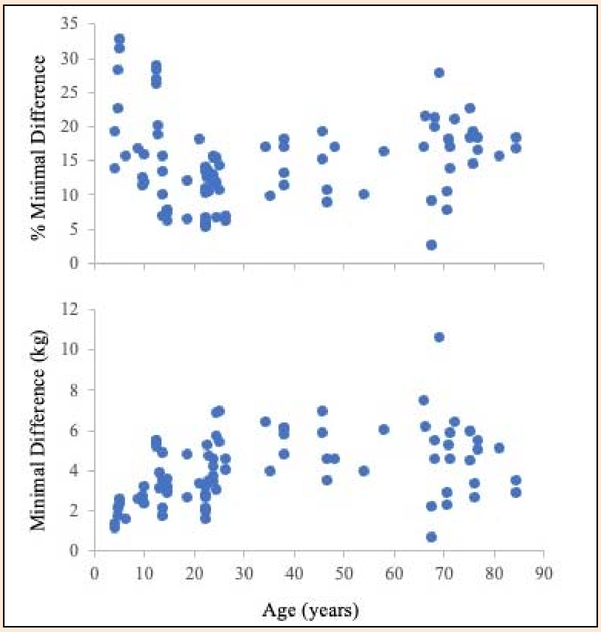

Figure 2. **The relationship between the average age of study participants and the absolute reliability [minimal difference (MD) and percentage of minimal difference to the measured value (%MD)] of handgrip strength measurements. In studies where MD was calculated using handgrip strength values of both hands, 1/2 MD value was used.** %MD = MD / average handgrip strength * 100

Impact of potential moderators on absolute test-retest reliability of grip strength measurements

Nine of the 48 studies did not report the mean age of participants, only an age range (Dugdale et al., 2019; Espana-Romero et al., 2010; Gasior et al., 2020; Plant et al., 2016; Sanchez-Delgado et al., 2015; Suzuki et al., 2019; Svensson et al., 2008; Walamies and Turjanmaa, 1993; Ward and Adams, 2007). There was considerable variation among studies reporting absolute test-retest reliability (MD and %MD) in each age group (Figure 2 and 4). Specifically, the mean MD value for young children (under 7 years old) was 1.9 kg (Abe et al., 2022; Amado-Pacheco et al., 2019; Cadenas-Sanchez et al., 2016; King-Dowling et al., 2024; Sanchez-Delgado et al., 2015; Svensson et al., 2008), while it was 2.5 kg for children aged 7 to 10 (Espana-Romero et al., 2010; Fernandez-Santos et al., 2016; Gerodimos, 2012; Gerodimos and Karatrantou, 2013; Gasior et al., 2020). The mean MD for adolescents (10 - 18 years old) was 4.2 kg (Dugdale et al., 2019; Espana-Romero et al., 2010; Gasior et al., 2020; Gerodimos, 2012; Gerodimos and Karatrantou, 2013; O'Keeffe et al., 2020; Ortega et al., 2008; Ramirez-Velez et al., 2015; Svensson et al., 2008; Trajkovic et al., 2024), which is similar to young adults (18 - 35 years old; 4.0 kg) (Balogun et al., 1991; Biasini et al., 2023; Boshnjaku et al., 2021; Cildan Uysal et al., 2022; Gerodimos, 2012; Gil et al., 2022; Hamilton et al., 1992; Karatrantou et al., 2020; Leszczak et al., 2024; Maurya et al., 2023; Petersen et al., 2015; Savva et al., 2013; Venegas-Carro et al., 2022). Middle-aged (36 - 60 years old) and older (>60 years old) adults had an MD of approximately 5 - 6 kg. On the other hand, the mean %MD values in young children and adolescents were approximately 25% and 17%, respectively, while those in young and older adults were about 12% and 17%, respectively (Table 4).

Mean weighted reliability statistics

The overall weighted MD was 4.463 (95% confidence interval [CI]: 3.926, 4.999; $p < 0.001$). As there was significant heterogeneity ($Q = 36{,}484.970$ $p < 0.001$) that could be attributed to both within (33.9%) and between (65.7%) study variance, potential moderators were examined. Neither age [effect size [ES]: 0.015 (95% CI: -0.004, 0.035; $p = 0.113$)], test-retest interval [ES: 0.006 (95% CI: -0.002, 0.013; $p = 0.143$)] nor handgrip device ($p = 0.752$) were significant moderators of reliability. The overall weighted %MD was 16.307 (95% CI: 14.529, 18.085; $p < 0.001$). Like that of the absolute MD, neither age [ES: -0.025 (95% CI: -0.089, 0.039; $p = 0.439$)], test-retest interval [ES: 0.022 (95% CI: -0.001, 0.046; $p = 0.065$)], or handgrip device ($p = 0.334$) were significant moderators of reliability. There was also no apparent systematic bias [ES: 0.162 (95% CI: -0.139, 0.464; $p = 0.291$)], and the presence of systematic bias was not moderated by age [ES: 0.005 (95%

CI: -0.006, 0.017; p = 0.380)] suggesting there was no learning effect, and this did not differ based on age.

Four studies reported absolute reliability for both boys and girls (Amado-Pacheco et al., 2019; Cadenas-Sanchez et al., 2016; Ortega et al., 2008; Ramirez-Velez et al., 2015; Sanchez-Delgado et al., 2015), while three studies focused on adult men and women (Bohannon, 2006; Karatrantou et al., 2020; Maurya et al., 2023). The mean MD value was 3.0 kg for boys and 2.5 kg for girls, with %MD values of 24.2% and 23.4%, respectively. In adults, the mean MD values were 3.9 kg for women and 5.4 kg for men, with %MD values of 17.3% for women and 15.3% for men (Table 4).

No studies have compared the absolute reliability of grip strength measurements between participants with and without sports experience. However, studies have been done on pre-pubertal, pubertal, and young adult basketball players (Gerodimos, 2012), pre-pubertal and pubertal wrestlers (Gerodimos and Karatrantou, 2013), youth soccer players (Dugdale et al., 2019), and middle-aged ten-pin bowlers (Tan et al., 2001). Moreover, twenty-two studies measured grip strength in both the left and right hands; half (11 studies) compared dominant and non-dominant hands, and the remaining 11 studies were able to compare right and left hands (Table 4).

Discussion

The current manuscript investigated the impact of potential moderating factors on the absolute test-retest reliability of grip strength measurements in a healthy population. This systematic review with meta-analysis included 48 studies involving 4,980 healthy participants (i.e., 2,025 children and adolescents, 1,451 young and middle-aged adults, and 1,504 older adults). Our findings demonstrated that (1) there was considerable variation among studies reporting MD and %MD across each age group; (2) the mean MD (%MD) values were 1.9 kg (25.4%) in young children (<7 years old), 2.5 kg (13.8%) in children (7 - 10 years old), 4.2 kg (17.1%) in adolescents (10 - 18 years old), 4.0 kg (11.6%) in young adults (18 - 35 years old), and 4.7 kg (16.7%) in older adults (>60 years old); (3) no studies have compared the MD and %MD between participants with and without sports experience; (4) neither age, test-retest interval, nor handgrip device served as significant moderators of MD and %MD reliability.

In this study, our meta-analysis found no evidence that the MD and %MD in test-retest reliability for grip strength measurements were influenced by age. One possible reason is the considerable variation in MD and %MD among studies within each age group (Figure 2 and Table 4). Nonetheless, the mean %MD for the reliability of test-retest grip strength measures in each age group is distinctive and partly similar to the %MD observed for muscle strength measures other than grip strength. For instance, maximal voluntary isokinetic muscle strength is a standard outcome measure for assessing knee joint function. A study measured knee extension and flexion peak torque at an angular velocity of 60 degrees per second across two sessions, 7 days apart, involving 22 children (10 boys and 12 girls) with a mean age of 8.8 years (Fagher et al., 2016). The MD and %MD calculated from the SEM of the test-retest reliability reported by the authors were 15.2 Nm and 30.9% for knee extension and 9.7 Nm and 36.1% for knee flexion, respectively. Another study (Santos et al., 2013) also assessed the test-retest reliability (7 days apart) of knee extension and flexion peak torque (60 degrees per second) in children with a mean age of 8.5 years. The %MD calculated from SEM was 29.3% for the dominant leg, 33.1% for the non-dominant leg for knee extension, and 46.2% and 32.6%, respectively, for knee flexion. The %MD values in these studies were similar to those observed in young children (< 7 years old) for grip strength measurements. Considering that the mean %MD of grip strength measurements in children of the same age group (7 - 10 years old) was 13.8%, the %MD of the isokinetic strength measure may appear high. In addition, a study (Maffiuletti et al., 2007) investigating the reproducibility (7 days apart) of knee extension and flexion peak torques under the same conditions (60 degrees per second) in young adults found that the %MD values are in the same range (10.7% for knee extension and 8.6% for knee flexion) as those observed in grip strength measurements in young adults (11.6%).

Eighty-five percent (41 studies) of the 48 included studies had a test-retest interval of less than two weeks, with 7 days being the most common (16 studies). This may explain why the test-retest interval did not affect grip strength measurements' MD and %MD. Those results suggest that at least a test-retest interval of two weeks or less may not significantly affect the grip strength reliability of MD and %MD. Three included studies reported test-retest reliability at two different intervals: 7 days vs. 9 weeks (Venegas-Carro et al., 2022), 24 hours vs. 1 year (Abe et al., 2018), and 12 weeks vs. 24 weeks (Jenkins and Cramer, 2017). For instance, Venegas-Carro et al. (2022) reported that MD and %MD values doubled at a 9-week interval (5.7 kg and 11.8%) compared with a 7-day interval (3.1 kg and 6.6 %). Abe et al. (2018) observed that although this test was performed on a different population, the MD and %MD values were greater at a 1-year interval (6.4 kg and 21.1 %) than at a 24-hour interval (3.97 kg and 10.1 %). However, Jenkins and Cramer (2017) reported similar MD values at 12- and 24-week intervals, making it unclear whether and at what point extending the interval affects grip strength reproducibility. Grip strength is one part of a physical fitness test taken annually for children and adolescents. Future studies may clarify the impact of extending the test-retest interval on the reproducibility of grip strength measurements.

About 70% of the included studies utilized the Jamar hand dynamometer, regarded as the gold standard or the Takei dynamometer. In both Jamar and Takei, despite differing standardized measurement conditions (sitting vs. standing, elbows at 90 degrees vs. extended, five grip widths vs. adjustments for hand size), the type of device did not affect MD and %MD in grip strength measurements. Furthermore, several studies examining the measurement accuracy of different handgrip dynamometers using Jamar as a benchmark also reported a good correlation between the two (Cildan Uysal et al., 2022; Hamilton et al., 1992; Trajkovic et al., 2024). However, the mechanical systems of the devices differ between Jamar (hydraulic) and Takei

(Smedley), and it has been observed that the measured values of the spring-type Takei differ from those of the Jamar in participants with high grip strength (Abe et al., 2019). Studies reporting test-retest reliability of grip strength measurement in young children are limited to studies using the Takei dynamometer.

In this study, sex and sports experience were not used as moderating variables in meta-analyses due to the limited number of studies that assessed boys and girls or men and women separately and the limited number of studies that assessed athletes. The difference in grip strength between boys and girls in children under 10 is less than that observed in adult men and women (Ramirez-Velez et al., 2021). Thus, MD may be similar in younger children when %MD is identical in both sexes. In adults, there is a clear sex difference in average grip strength, and when the %MD is the same for both sexes, it is expected that the MD of women with lower grip strength will be smaller than that of men. The results of included studies reporting MD and %MD separately for boys and girls or men and women suggest this possibility (Table 4), although future studies are needed. In addition, no studies have compared MD and %MD between athletes and non-athletes. Further research may be required to clarify whether sports experience affects the test-retest reliability of grip strength measurements.

Conclusion

The data analyzed from the collected studies found considerable variation among studies reporting MD and the percentage of MD to measured value (%MD) across each age group. Neither age, test-retest interval nor handgrip device served as significant moderators of MD and %MD reliability. Due to the limited number of studies, sex and sports experience were excluded from the analysis; as a result, their impacts remain unknown.

Acknowledgements

The authors have no conflict of interest to declare. This study received no specific grants, fellowships, or materials gifts from any funding agency in the public, commercial, or not-for-profit sectors. The datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author who organized the study.

References

Abe, T., Thiebaud, R. S. and Loenneke, J. P. (2016) Age-related change in handgrip strength in men and women: Is muscle quality a contributing factor? *Age* **38(1)**, 28. https://doi.org/10.1007/s11357-016-9891-4

Abe, T., Dankel, S. J., Buckner, S. L., Jessee, M. B., Mattocks, K. T., Mouser, J. G., Bell, Z. W. and Loenneke, J. P. (2018) Short term (24 hours) and long term (1 year) assessments of reliability in older adults: Can one replace the other? *Journal of Aging Research and Clinical Practice* **7**, 82-84. https://doi.org/10.14283/jarcp.2018.15

Abe, T., Loenneke, J. P., Thiebaud, R. S. and Loftin, M. (2019) The bigger the hand, the bigger the difference? Implications for testing strength with 2 popular handgrip dynamometers. *Journal of Sport Rehabilitation* **28(3)**, 278-282. https://doi.org/10.1123/jsr.2017-0189

Abe, T., Thiebaud, R. S., Ozaki, H., Yamasaki, S. and Loenneke, J. P. (2022) Children with low handgrip strength: A narrative review of possible exercise strategies to improve its development. *Children (Basel)* **9(11)**, 1616. https://doi.org/10.3390/children9111616

Abe, T., Sanui, R., Sasaki, A., Ishibashi, A., Daikai, N., Shindo, Y., Abe, A. and Loenneke, J. P. (2022) Optimal grip span for measuring maximum handgrip strength in preschool children. *International Journal of Clinical Medicine* **13(11)**, 479-488. https://doi.org/10.4236/ijcm.2022.1311035

Abe, T., Machida, S., Nakamura, M., Kohmura, Y., Suzuki, K., Abe, A., Nakano, M., Loenneke, J. P. and Naito, H. (2024) Tracking handgrip strength in Kendo athletes from university to middle and older adulthood. *American Journal of Human Biology* **36(9)**, e24082. https://doi.org/10.1002/ajhb.24082

Amado-Pacheco, J. C., Prieto-Benavides, D. H., Correa-Bautista, J. E., Garcia-Hermoso, A., Agostinis-Sobrinho, C., Alonso-Martinez, A. M., Izquierdo, M. and Ramirez-Velez, R. (2019) Feasibility and reliability of physical fitness tests among Colombian preschool children. *International Journal of Environmental Research and Public Health* **16(17)**, 3069. https://doi.org/10.3390/ijerph16173069

Anstey, K. J., Smith, G. A. and Lord, S. (1997) Test-retest reliability of a battery of sensory, motor and physiological measures of aging. *Perceptual and Motor Skills* **84(3 Pt 1)**, 831-834. https://doi.org/10.2466/pms.1997.84.3.831

Assink, M. and Wibbelink, C. J. (2016) Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology* **12(3)**, 154-174. https://doi.org/10.20982/tqmp.12.3.p154

Balogun, J. A., Adenlola, S. A. and Akinloye, A. A. (1991) Grip strength normative data for the Harpenden dynamometer. *Journal of Orthopedic Sports Physical Therapy* **14(4)**, 155-160. https://doi.org/10.2519/jospt.1991.14.4.155

Beauchamp, M. K., Hao, Q., Kuspinar, A., D'Amore, C., Scime, G., Ma, J., Mayhew, A., Bassim, C., Wolfson, C., Kirkland, S., Griffith, L. and Raina, P. (2021) Reliability and minimal detectable change values for performance-based measures of physical functioning in Canadian Longitudinal Study on Aging. *Journal of Gerontology A Biological Science and Medical Science* **76(11)**, 2030-2038. https://doi.org/10.1093/gerona/glab175

Biasini, N. R., Pellegrino, M., Switzer-McIntyre, S. and Kasawara, K. T. (2023) Reliability and validity of shoulder and handgrip strength testing. *Physiotherapy Canada* **75(1)**, 65-71. https://doi.org/10.3138/ptc-2021-0041

Bobos, P., Nazari, G., Lu, S. and MacDermid, J. C. (2020) Measurement properties of the hand grip strength assessment: A systematic review with meta-analysis. *Archives of Physical Medicine and Rehabilitation* **101(3)**, 553-565. https://doi.org/10.1016/j.apmr.2019.10.183

Bohannon, R. W. (2015) Muscle strength: Clinical and prognostic value of hand-grip dynamometry. *Current Opinion in Clinical Nutrition & Metabolic Care* **18(5)**, 465-470. https://doi.org/10.1097/MCO.0000000000000202

Bohannon, R. W. (2017) Test-retest reliability of measurements of hand-grip strength obtained by dynamometry from older adults: A systematic review of research in the PubMed Database. *Journal of Frailty & Aging* **6(2)**, 83-87. https://doi.org/10.14283/jfa.2017.8

Bohannon, R. W. (2019) Grip strength: An indispensable biomarker for older adults. *Clinical Intervention in Aging* **14**, 1681-1691. https://doi.org/10.2147/CIA.S194543

Bohannon, R. W. (2006) Test-retest reliability of the MicroFET 4 hand-grip dynamometer. *Physiotherapy Theory and Practice* **22(4)**, 219-221. https://doi.org/10.1080/09593980600822875

Bohannon, R. W. and Schaubert, K. L. (2005) Test-retest reliability of grip-strength measures obtained over a 12-week interval from community-dwelling elders. *Journal of Hand Therapy* **18(4)**, 426-427. https://doi.org/10.1197/j.jht.2005.07.003

Bohannon, R. W., Bubela, D. J., Magasi, S. and Gershon, R. C. (2011) Relative reliability of three objective tests of limb muscle strength. *Isokinetic and Exercise Science* **19(2)**, 77-81. https://doi.org/10.3233/IES-2011-0400

Boshnjaku, A., Bahtiri, A., Feka, K., Krasniqi, E., Tschan, H. and Wessner, B. (2021) Test-retest reliability data of functional performance, strength, peak torque and body composition assessments in two different age groups of Kosovan adults. *Data in Brief* **36**, 106988. https://doi.org/10.1016/j.dib.2021.106988

Brink, Y. and Louw, Q. A. (2011) Clinical instruments: Reliability and validity critical appraisal. *Journal of Evaluation in Clinical*

Practice **18(6)**, 1126-1132. https://doi.org/10.1111/j.1365-2753.2011.01707.x

Cadenas-Sanchez, C., Martinez-Tellez, B., Sanchez-Delgado, G., Mora-Gonzalez, J., Castro-Pinero, J., Lof, M., Ruiz, J. R. and Ortega, F. B. (2016) Assessing physical fitness test in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *Journal of Science and Medicine in Sport* **19(11)**, 910-915. https://doi.org/10.1016/j.jsams.2016.02.003

Celis-Morales, C. A., Lyall, D. M., Anderson, J., Iliodromiti, S., Fan, Y., Ntuk, U. E., Mackay, D. F., Pell, J. P., Sattar, N. and Gill, J. M. R. (2017) The association between physical activity and risk of mortality is modulated by grip strength and cardiorespiratory fitness: Evidence from 498135 UK-Biobank participants. *European Heart Journal* **38(2)**, 116-122. https://doi.org/10.1093/eurheartj/ehw249

Cildan Uysal, S., Tonak, H. A. and Kitis, A. (2022) Validity, reliability and test-retest study of grip strength measurement in two positions with two dynamometers: Jamar® Plus and K-Force® Grip. *Hand Surgery and Rehabilitation* **41(3)**, 305-310. https://doi.org/10.1016/j.hansur.2022.02.007

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boitie, Y., Bruyere, O., Cederholm, T., Cooper, C., Landi, F., Rolland, Y., Sayer, A. A., Schneider, S. M., Sieber, C. C., Topinkova, E., Vandewoude, M., Visser, M. and Zamboni, M. (2019) Sarcopenia: Revised European consensus on definition and diagnosis. *Age and Ageing* **48(4)**, 16-31. https://doi.org/10.1093/ageing/afz046

Dankel, S. J., Kang, M., Abe, T. and Loenneke, J. P. (2019) A meta-analysis to determine the validity of taking blood pressure using the indirect cuff method. *Current Hypertension Reports* **21(1)**, 11. https://doi.org/10.1007/s11906-019-0929-8

Dugdale, J. H., Arthur, C. A., Sanders, D. and Hunter, A. M. (2019) Reliability and validity of field-based fitness tests in youth soccer players. *European Journal of Sport Science* **19(6)**, 745-756. https://doi.org/10.1080/17461391.2018.1556739

Espana-Romero, V., Artero, E. G., Jimenez-Pavon, D., Cuenca-Garcia, M., Ortega, F. B., Castro-Pinero, J., Sjostrom, M., Castillo-Garzon, M. J. and Ruiz, J. R. (2010) Assessing health-related fitness tests in the school setting: Reliability, feasibility and safety; The ALPHA Study. *International Journal of Sports Medicine* **31(7)**, 490-497. https://doi.org/10.1055/s-0030-1251990

Essendrop, M., Schibye, B. and Hansen, K. (2001) Reliability of isometric muscle strength tests for the trunk, hands and shoulders. *International Journal of Industrial Engineering* **28(6)**, 379-387. https://doi.org/10.1016/S0169-8141(01)00044-0

Fagher, K., Fritzson, A. and Drake, A. M. (2016) Test-retest reliability of isokinetic knee strength measurements in children aged 8 to 10 years. *Sports Health* **8(3)**, 255-259. https://doi.org/10.1177/1941738116632506

Fernandez-Santos, J. R., Ruiz, J. R., Gonzalez-Montesinos, J. L. and Castro-Pinero, J. (2016) Reliability and validity of field-based tests to assess upper-body muscular strength in children aged 6-12 years. *Pediatric Exercise Science* **28(2)**, 331-340. https://doi.org/10.1123/pes.2015-0196

Ferreira, S., Raimundo, A. and Marmeleira, J. (2021) Test-retest reliability of the functional reach test and the hand grip strength test in older adults using nursing home services. *Irish Journal of Medical Science* **190(4)**, 1625-1632. https://doi.org/10.1007/s11845-020-02492-0

Gasior, J. S., Pawlowski, M., Jelen, P. J., Rameckers, E. A., Williams, C. A., Makuch, R. and Werner, B. (2020) Test-retest reliability of handgrip strength measurement in children and preadolescents. *International Journal of Environmental Research and Public Health* **17(21)**, 8026. https://doi.org/10.3390/ijerph17218026

Gerodimos, V. (2012) Reliability of handgrip strength test in basketball players. *Journal of Human Kinetics* **31**, 25-36. https://doi.org/10.2478/v10078-012-0003-y

Gerodimos, V. and Karatrantou, K. (2013) Reliability of maximal handgrip strength test in pre-pubertal and pubertal wrestlers. *Pediatric Exercise Science* **25(2)**, 308-322. https://doi.org/10.1123/pes.25.2.308

Gil, A. W., da Silva, R. A., Pereira, C., Nascimento, V. B., Amorim, C. F., Imaizumi, M. and Teixeira, D. C. (2022) Reproducibility of dynamometers in handrail format in evaluating handgrip strength and traction in young and older adults. *Medical Engineering & Physics* **100**, 103749. https://doi.org/10.1016/j.medengphy.2021.103749

Hamilton, G. F., McDonald, C. and Chenier, T. C. (1992) Measurement of grip strength: Validity and reliability of the sphygmomanometer and Jamar grip dynamometer. *Journal of Orthopaedic & Sports Physical Therapy* **16(5)**, 215-219. https://doi.org/10.2519/jospt.1992.16.5.215

Hanten, W. P., Chen, W. Y., Austin, A. A., Brooks, R. E., Carter, H. C., Law, C. A., Morgan, M. K., Sanders, D. J., Swan, C. A. and Vanderslice, A. L. (1999) Maximum grip strength in normal subjects from 20 to 64 years of age. *Journal of Hand Therapy* **12(3)**, 193-200. https://doi.org/10.1016/s0894-1130(99)80046-5

Jenkins, N. D. M. and Cramer, J. T. (2017) Reliability and minimum detectable change for common clinical physical function tests in sarcopenic men and women. *Journal of American Geriatrics Society* **65(4)**, 839-846. https://doi.org/10.1111/jgs.14769

Karatrantou, K., Xagorari, A., Vasilopoulou, T. and Gerodimos, V. (2020) Does the number of trials affect the reliability of handgrip strength measurement in individuals with intellectual disability? *Hand Surgery and Rehabilitation* **39(3)**, 223-228. https://doi.org/10.1016/j.hansur.2020.01.004

Kieser, J., Langford, M., Stover, E., Tomkinson, G. R., Clark, B. C., Cawthon, P. M. and McGrath, R. (2025) Absolute agreement between subjective hand squeeze and objective handgrip strength in adults. *Journal of Strength and Conditioning Research* **39(1)**, 16-23. https://doi.org/10.1519/JSC.0000000000004936

King-Dowling, S., Fortnum, K., Chirico, D., Le, T., Kwan, M. Y. W., Timmons, B. W. and Cairney, J. (2024) Reliability of field- and laboratory-based assessments of health-related fitness in preschool-aged children. *American Journal of Human Biology* **36(2)**, e23987. https://doi.org/10.1002/ajhb.23987

Legg, H. S., Spindor, J., Dziendzielowski, R., Sharkey, S., Lanovaz, J. L., Farthing, J. P. and Arnold, C. M. (2020) The reliability and validity of novel clinical strength measures of the upper body in older adults. *Hand Therapy* **25(4)**, 130-138. https://doi.org/10.1177/1758998320957373

Lemmink, K. A. P. M., Han, K., de Greef, M. H. G., Rispens, P. and Stevens, M. (2001) Reliability of the Groningen fitness test for the elderly. *Journal of Aging and Physical Activity* **9(2)**, 194-212. https://doi.org/10.1123/japa.9.2.194

Leszczak, J., Pniak, B., Druzbicki, M. and Guzik, A. (2024) The reliability of a biometrics device as a tool for assessing hand grip and pinch strength, in a Polish cohort - A prospective observational study. *PLoS One* **19(5)**, e0303648. https://doi.org/10.1371/journal.pone.0303648

Loenneke, J. P., Abe, A., Yamasaki, S., Tahara, Y. and Abe, T. (2024) Sex differences in strength during development: Implications for inclusivity and fairness in sport. *American Journal of Human Biology* **36(11)**, e24152. https://doi.org/10.1002/ajhb.24152

Maffiuletti, N. A., Bizzini, M., Desbrosses, K., Babault, N. and Munzinger, U. (2007) Reliability of knee extension and flexion measurements using the Com-Trex isokinetic dynamometer. *Clinical Physiology and Functional Imaging* **27(6)**, 346-353. https://doi.org/10.1111/j.1475-097X.2007.00758.x

Maurya, P. S., Sisneros, K. P., Johnson, E. B. and Palmer, T. B. (2023) Reliability of handgrip strength measurements and their relationship with muscle power. *Journal of Sports Medicine and Physical Fitness* **63(7)**, 805-811. https://doi.org/10.23736/S0022-4707.23.14820-1

Norman, K., Stobaus, N., Gonzalez, M. C., Schulzke, J. D. and Pirlich, M. (2011) Hand grip strength: Outcome predictor and marker of nutritional status. *Clinical Nutrition* **30(2)**, 135-142. https://doi.org/10.1016/j.clnu.2010.09.010

O'Keeffe, B. T., Donnelly, A. E. and MacDonncha, C. (2020) Test-retest reliability of student-administered health-related fitness tests in school settings. *Pediatric Exercise Science* **32(1)**, 48-57. https://doi.org/10.1123/pes.2019-0166

Ortega, F. B., Artero, E. G., Ruiz, J. R., Vicente-Rodriguez, G., Bergman, P., Hagstromer, M., Ottevaere, C., Nagy, E., Konsta, O., Rey-Lopez, J. P., Polito, A., Dietrich, S., Plada, M., Beghin, L., Manios, Y., Sjostrom, M. and Castillo, M. J. (2008) Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *International Journal of Obesity (Lond)* **32 Suppl 5**, S49-S57. https://doi.org/10.1038/ijo.2008.183

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M. and Moher, D. (2021) Updating guidance for reporting systematic reviews:

Development of the PRISMA 2020 statement. *Journal of Clinical Epidemiology* **134**, 103-112. https://doi.org/10.1016/j.jclinepi.2021.02.003
Peralta, M., Dias, C. M., Marques, A., Henriques-Neto, D. and Sousa-Uva, M. (2023) Longitudinal association between grip strength and the risk of heart diseases among European middle-aged and older adults. *Experimental Gerontology* **171**, 112014. https://doi.org/10.1016/j.exger.2022.112014
Petersen, N., Thieschafer, L., Ploutz-Snyder, L., Damann, V. and Mester, J. (2015) Reliability of a new test battery for fitness assessment of the European Astronaut corps. *Extreme Physiology & Medicine* **4**, 12. https://doi.org/10.1186/s13728-015-0032-y
Plant, C. E., Parsons, N. R., Edwards, A., Rice, H., Denninson, K. and Costa, M. L. (2016) A comparison of electronic and manual dynamometry and goniometry in patients with fracture of the distal radius and healthy participants. *Journal of Hand Therapy* **29(1)**, 73-80. https://doi.org/10.1016/j.jht.2015.11.004
Ramirez-Velez, R., Rincon-Pabon, D., Correa-Bautista, J. E., Garcia-Hermoso, A. and Izquierdo, M. (2021) Handgrip strength: Normative reference values in males and females aged 6-64 years old in a Colombian population. *Clinical Nutrition ESPEN* **44**, 379-386. https://doi.org/10.1016/j.clnesp.2021.05.009
Ramirez-Velez, R., Rodrigues-Bezerra, D., Correa-Bautista, J. E., Izquierdo, M. and Lobelo, F. (2015) Reliability of health-related physical fitness tests among Colombian children and adolescents: The FUPRECOL study. *PLoS One* **10(10)**, e014875. https://doi.org/10.1371/journal.pone.0140875
Rantanen, T., Volpato, S., Ferrucci, L., Heikkinen, E., Fried, L. P. and Guralnik, J. M. (2003) Handgrip strength and cause-specific and total mortality in older disabled women: Exploring the mechanism. *Journal of American Geriatrics Society* **51(5)**, 636-641. https://doi.org/10.1034/j.1600-0579.2003.00207.x
Saint-Maurice, P. F., Laurson, K., Welk, G. J., Eisenmann, J., Gracia-Marco, L., Artero, E. G., Ortega, F., Ruiz, J. R., Moreno, L. A., Vicente-Rodriguez, G. and Janz, K. F. (2018) Grip strength cutpoints for youth based on a clinically relevant bone health outcome. *Archives of Osteoporosis* **13(1)**, 92. https://doi.org/10.1007/s11657-018-0502-0
Sanchez-Delgado, G., Cadenas-Sanchez, C., Mora-Gonzalez, J., Martinez-Tellez, B., Chillon, P., Lof, M., Ortega, F. B. and Ruiz, J. R. (2015) Assessment of handgrip strength in preschool children aged 3 to 5 years. *Journal of Hand Surgery* **40(9)**, 966-972. https://doi.org/10.1177/1753193415592328
Santos, A. N., Pavao, S. L., Avila, M. A., Salvini, T. F. and Rocha, N. A. C. F. (2013) Reliability of isokinetic evaluation in passive mode for knee flexors and extensors in healthy children. *Brazilian Journal of Physical Therapy* **17(2)**, 112-120. https://doi.org/10.1590/S1413-35552012005000074
Savva, C., Karagiannis, C. and Rushton, A. (2013) Test-retest reliability of grip strength measurement in full elbow extension to evaluate maximum grip strength. *Journal of Hand Therapy* **38(2)**, 183-186. https://doi.org/10.1177/1753193412449804
Stenholm, S., Tiainen, K., Rantanen, T., Sainio, P., Heliovaara, M., Impivaara, O. and Koskinen, S. (2012) Long-term determinants of muscle strength decline: Prospective evidence from the 22-year Mini-Finland follow-up survey. *Journal of American Geriatrics Society* **60(1)**, 77-85. https://doi.org/10.1111/j.1532-5415.2011.03779.x
Suzuki, Y., Kamide, N., Kitai, Y., Ando, M., Sato, H., Yoshitake, S. and Sakamoto, M. (2019) Absolute reliability of measurements of muscle strength and physical performance measures in older people with high functional capacities. *European Geriatrics Medicine* **10(5)**, 733-740. https://doi.org/10.1007/s41999-019-00218-9
Svensson, E., Waling, K. and Hager-Ross, C. (2008) Grip strength in children: Test-retest reliability using Grippit. *Acta Paediatrica* **97(9)**, 1226-1231. https://doi.org/10.1111/j.1651-2227.2008.00895.x
Tan, B., Aziz, A. R., Teh, K. C. and Lee, H. C. (2001) Grip strength measurement in competitive ten-pin bowlers. *Journal of Sports Medicine Physical Fitness* **41(1)**, 68-72.
Trajkovic, N., Rancic, D., Ilic, T., Herodek, R., Korobeynikov, G. and Pekas, D. (2024) Measuring handgrip strength in school children: Inter-instrument reliability between Takei and Jamar. *Scientific Reports* **14(1)**, 1074. https://doi.org/10.1038/s41598-024-51368-1
Tsang, R. C. C. (2005) Reference values for 6-minute walk test and hand-grip strength in healthy Hong Kong Chinese adults. *Hong Kong Physiotherapy Journal* **23(1)**, 6-12. https://doi.org/10.1016/S1013-7025(09)70053-3
Venegas-Carro, M., Kramer, A., Moreno-Villanueva, M. and Gruber, M. (2022) Test-retest reliability and sensitivity of common strength and power tests over a period of 9 weeks. *Sports* **10(11)**, 171. https://doi.org/10.3390/sports10110171
Villafane, J. H., Valdes, K., Buraschi, R., Martinelli, M., Bissolotti, L. and Negrini, S. (2016) Reliability of the handgrip strength test in elderly subjects with Parkinson disease. *Hand* **11(1)**, 54-58. https://doi.org/10.1177/1558944715614852
Walamies, M. and Turjanmaa, V. (1993) Assessment of the reproducibility of strength and endurance handgrip parameters using a digital analyser. *European Journal of Applied Physiology* **67(1)**, 83-86. https://doi.org/10.1007/BF00377710
Ward, C. and Adams, J. (2007) Comparative study of the test-re-test reliability of four instruments to measure grip strength in a healthy population. *British Journal of Hand Therapy* **12(2)**, 48-54. https://doi.org/10.1177/175899830701200202
Weir, J. P. (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research* **19(1)**, 231-240. https://doi.org/10.1519/15184.1

Key points

- In previous studies, the intraclass correlation coefficient (ICC) is often the preferred method for reporting measurement reliability.
- One major limitation with reporting ICC values is that they are entirely dependent on the heterogeneity of the sample included in the reliability assessment.
- While there are certainly instances where relative reliability may be necessary, often absolute reliability is preferred and more useful.
- We found considerable variation among studies reporting absolute test-retest reliability of grip strength tests, such as minimal differences (MD) and the percentage of MD to the measured value across each age group.
- Neither age, test-retest interval, nor handgrip device served as a significant moderator of MD and percentage of MD reliability.

AUTHOR BIOGRAPHY

Takashi ABE
Employment
Institute of Health and Sports Science & Medicine, Juntendo University
Degree
PhD
Research interests
Exercise physiology and adaptation by exercise training
E-mail: t12abe@gmail.com

Jun Seob SONG
Employment
Department of Counseling, Health and Kinesiology, Texas A&M University-San Antonio
Degree
PhD
Research interests
Skeletal muscle physiology
E-mail: jsong1@tamusa.edu

Scott J. DANKEL
Employment
Department of Health and Exercise Science, Rowan University
Degree
PhD
Research interests
Exercise physiology and statistics
E-mail: dankel47@rowan.edu

Ricardo B. VIANA
Employment
Institute of Physical Education and Sport, Federal University of Ceará
Degree
PhD
Research interests
Exercise physiology and human anatomy
E-mail: vianaricardoborges@ufc.br

Akemi ABE
Employment
Division of Children's Health and Exercise Research, Institute of Trainology
Degree
BS
Research interests
Active play and health in children
E-mail: amyabe3379@gmail.com

Jeremy P. LOENNEKE
Employment
Department of Health, Exercise Science, and Recreation Management, The University of Mississippi
Degree
PhD
Research interests
Skeletal muscle physiology
E-mail: jploenne@olemiss.edu

✉ **Takashi Abe, PhD**
Institute of Health and Sports Science & Medicine, Juntendo University, 1-1 Hirakagakuendai, Inza-shi, Chiba 270-1695, Japan