J. Phys. Ther. Sci. 36: 598–602, 2024

The Journal of Physical Therapy Science

Original Article

Test-retest reliability of the ten-repetition maximum test in untrained young males and females

Masahito Taga, MS[1, 2)*], Hotaka Sugimoto, MS[2, 3)], Kimito Momose, PhD[4)]

[1)] *Department of Rehabilitation, Ina Central Hospital: 1313-1 Koshiroukubo, Ina-shi, Nagano 396-8555, Japan*
[2)] *Department of Health Sciences, Graduate School of Medicine, Shinshu University, Japan*
[3)] *Division of Physical Therapy, Department of Rehabilitation, Faculty of Health Sciences, Nagano University of Health and Medicine, Japan*
[4)] *Department of Physical Therapy, School of Health Science, Shinshu University, Japan*

Abstract. [Purpose] This study investigated the test-retest reliability of the ten-repetition maximum leg press (LP) and chest press (CP) tests in untrained young males and females. The secondary aim was to examine the test-retest reliability of the 10RM test according to gender. [Participants and Methods] All participants underwent the ten-repetition maximum test for the LP and CP across three sessions separated by 2 to 6 days of recovery: the first session was familiarization, the second for the initial test, and the third for the retest. To evaluate test-retest reliability, we calculated the intraclass correlation coefficient, standard error of measurement, and minimal detectable change for both the LP and CP. These measures were analyzed separately for each gender to assess gender-related differences. [Results] High intraclass correlation coefficient index was obtained for both LP and CP for both genderes. The standard errors of measurement and minimal detectable changes were consistent with previously reported values. However, females showed more variability than males, particularly in the LP. [Conclusion] The ten-repetition maximum test for LP and CP is a reliable method for measuring strength in untrained young males and females. However, it should be noted that the margin of error may be higher among female participants compared to their male counterparts.

Key words: Strength measurement, Ten-repetition maximum test, Reliability

(This article was submitted May 24, 2024, and was accepted Jun. 29, 2024)

INTRODUCTION

Muscle weakness causes fractures, falls, and a decline in physical performance[1–3)]. Resistance training is often used to prevent muscle weakness. Resistance training increases muscle strength, muscle size, and local muscular endurance[4)]. Furthermore, resistance training affects various aspects of the human body, such as mortality, disease incidence, and cognitive function, including the metabolic system[5, 6)]. Resistance training is a common program in physical therapy clinical settings, and to maximize the training effect, accurately evaluating individual strength is warranted.

The gold standard for measuring dynamic muscle strength is one-repetition maximum (1RM) test[7)]. However, the 1RM test induces muscle soreness and severe muscle damage, especially in beginners[8)]. As the 1RM test involves handling very high weights, the risk of acute injury is high[9)]. To avoid the risk of injury associated with the 1RM test, research has been conducted to predict the 1RM using the multiple-RM test[8, 10–12)].

*Corresponding author. Masahito Taga (E-mail: armed.t.rex@gmail.com)

Multiple-RM tests, such as the 10RM test, have a lower maximum weight and may have a lower risk of measurement-related injury than the 1RM test. If the multiple-RM test can accurately measure an individual's muscle strength, evaluating an individual's muscle strength than the 1RM test would be safer. However, limited studies have examined the test-retest reliability of the multiple RM test. Thus, this study primarily aimed to investigate the test-retest reliability of the 10RM test in untrained young individuals by calculating the minimal detectable change (MDC) present in measurements. Additionally, by comparing the amount of MDC of the 1RM test with that reported in previous studies, we considered whether the 10RM test could be used as an alternative to the 1RM test. The second purpose was to investigate test-retest reliability by gender as previous research suggested that reproducibility may vary by gender[13].

PARTICIPANTS AND METHODS

To examine the test-retest reliability of the 10RM test, a cross-sectional study was conducted. At least 26 participants were required based on a preliminary power test with an effect size correlation of 0.50, alpha level of 0.05, and power of 0.80. Facility recruitment was conducted among local hospital staff and included healthy adults aged 20–35 years. Participants had not undergone any resistance training in a year prior to participation. The exclusion criteria were as follows: (a) restriction of motor activity by a physician; (b) presence of pain, possibly requiring the discontinuation of motor tasks; and (c) a history of orthopedic, neurological, or internal medical diseases that may impede training.

The study protocol was approved by the Ethics Committee of Shinshu University (approval number: 5318). Prior to the study, a written informed consent was obtained from all participants. This study complied with the ethical standards of the 1964 Declaration of Helsinki and its subsequent revisions.

All participants were required to visit our laboratory for 4 days. The first day was an explanation of the study and practice of the movement of the LP and CP with the equipment used. The equipment was a chest press machine and leg press machine, and the load was adjusted in 5-kg increments. In practice, they performed four sets of 10 repetitions with a very light load. On the second to fourth days, they conducted a 10RM test for the LP and CP. All measurements were performed by the same examiner. The second day was considered the familiarization session. We used data from the third (initial) and fourth (retest) days for the analysis. Each visit was separated by 2–6 days to avoid the effects of fatigue.

The 10RM test was conducted according to the National Strength and Conditioning Association procedures[14]. For warm-up, participants completed 5 min of ergometer cycling and stretching of the muscle groups to be used in LP and CP, and 10 repetitions using a light load, moderate load, and maximum load close to 10RM with a 1–2-min rest. The participants then attempt at 10RM. If the participants succeeded in an attempt, a load was added (5–10 kg); if they failed, the load was removed (5 kg). The rest time between each attempt was 3 min. The maximum load at which they could complete 10 repetitions was recorded as 10RM. These attempts at 10RM were completed in five sets or fewer. Exercises were performed in the following order: CP and LP.

Means and standard deviations (SD) were used to describe participant characteristics and 10RM scores. To evaluate test-retest reliability, we used the following statistical approaches: (a) intraclass correlation coefficient ([ICC]; model: two-way mixed effects, absolute agreement, single rater); (b) the standard error of measurement ([SEM]; SEM=SD of the test scores multiplied by the square root of 1 −ICC[15]; and SEM%); (c) minimal detectable change ([MDC]; MDC=1.96 × the square root of 2 × SEM[15]; and MDC%); (d) the Bland–Altman plots to evaluate the agreement between test-retest. To evaluate the gender differences, we conducted the aforementioned steps for each gender. All statistical analyses were performed using RStudio 2023.09.0+463 software (Posit Software, PBC, Boston, MA, USA).

RESULTS

This study required participation for 4 days; there were no dropouts or missing data. In this study, 11 males and 19 females were included. The participant characteristics are presented in Table 1 and the results of the agreement analysis are presented in Table 2. The results of Bland–Altman plots were shown in Fig. 1. Both LP and CP demonstrated a high ICC

Table 1. Participants' characteristic and 10RM values

		All participant (n=30)	Male (n=11)	Female (n=19)
Age (years)		25.50 ± 2.21	25.91 ± 1.22	25.26 ± 2.62
BMI (kg/m^2)		20.50 ± 2.31	20.98 ± 1.53	20.22 ± 2.66
10RM LP (kg)	Initial test	156.5 ± 54.5	210.0 ± 37.9	125.5 ± 35.1
	Retest	161.5 ± 54.9	212.3 ± 39.3	132.1 ± 38.8
10RM CP (kg)	Initial test	49.2 ± 21.4	73.6 ± 13.7	35.0 ± 7.5
	Retest	48.0 ± 21.1	72.3 ± 13.7	33.9 ± 6.6

Mean ± standard deviation. BMI: body mass index; RM: repetition maximum; LP: leg press; CP: chest press.

Table 2. Results of agreement analysis

		All participant (n=30)	Male (n=11)	Female (n=19)
Mean difference between tests (95% CI) (kg)	LP	5.00 (1.90 to 8.10)	2.27 (−1.80 to 6.35)	6.58 (2.18 to 10.98)
	CP	−1.17 (−1.97 to −0.36)	−1.36 (−2.93 to 0.21)	−1.05 (−2.06 to −0.04)
ICC (95% CI)	LP	0.98 (0.95 to 0.99)	0.99 (0.95 to 1.0)	0.96 (0.82 to 0.99)
	CP	0.99 (0.98 to 1.0)	0.98 (0.92 to 1.0)	0.95 (0.85 to 0.98)
SEM (SEM%) (kg)	LP	7.74 (4.87)	3.86 (1.22)	7.40 (5.74)
	CP	2.12 (4.37)	1.94 (2.66)	1.57 (4.56)
MDC (MDC%) (kg)	LP	21.45 (13.49)	10.71 (3.39)	20.51 (15.92)
	CP	5.89 (12.15)	5.38 (7.38)	4.36 (12.64)
LOALB (95% CI) (kg)	LP	−11.28 (−16.65 to −5.91)	−9.62 (−16.68 to −2.56)	−11.33 (−18.96 to −3.70)
	CP	−5.38 (−6.77 to −3.99)	−5.94 (−8.66 to −3.22)	−5.16 (−6.91 to −3.41)
LOAUB (95% CI) (kg)	LP	21.28 (15.91 to 26.65)	14.17 (7.11 to 21.22)	24.49 (16.86 to 32.11)
	CP	3.05 (1.66 to 4.44)	3.21 (0.50 to 5.93)	3.05 (1.30 to 4.80)
Correlation analysis	LP	ns	ns	ns
	CP	ns	ns	ns

CI: confidence interval; ICC: intraclass correlation coefficient; SEM: standard error of measurement; MDC: minimal detectable change; LOALB: limits of agreement lower boundary; LOAUB: limits of agreement upper boundary; LP: leg press; CP: chest press; ns: not significant.

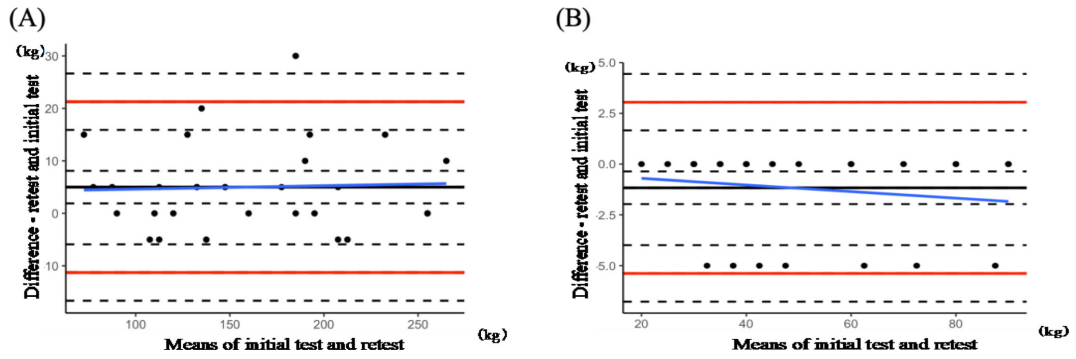

Fig. 1. Bland–Altman graphical representation of the comparison between initial test and retest. (A) and (B) represent the agreement between initial test and retest for LP and CP, respectively.
The solid black line indicates the mean of the differences between tests.
The solid red lines indicate limits of agreement for the upper and lower boundary.
The solid blue line indicates regression between tests.
Dashed lines show the 95% confidence interval of the means of differences and limits of agreement.

index (LP=0.98, CP=0.99), even when divided by gender (male: LP=0.99, CP=0.98; female: LP=0.96, CP=0.95). In LP, the difference of test-retest was 5.00 kg (± 3.10), absolute SEM of 7.74 kg, and absolute MDC of 21.45 kg. In CP, the difference of the test-retest was −1.17 kg (± 0.80), absolute SEM of 2.12 kg, and absolute MDC of 5.89 kg. While the retest was higher than the initial test in the LP, it was lower than the initial test in the CP. As a result of the analysis by gender, in female LP, the SEM and MDC were twice as large as in males.

DISCUSSION

This study aimed to investigate whether there is test-retest agreement of the 10RM test for LP and CP in untrained young males and females. The results show a high ICC index, indicating high reproducibility. Dias Fernandes et al. reported absolute SEM of 1.73 kg and absolute MDC of 4.79 kg for CP[15]. The study results are comparable to those of a previous study, which demonstrated a small standard error of measurement and a minimal detectable change (SEM of 2.12 kg, MDC 5.89 kg). The %MDC was higher than approximately 7 kg in the study by Dias Fernandes et al., which indicated a difference in the average strength of the participants (mean value was 24.96 kg in Dias Fernandes's study, 49.2 kg in this study).

A comparison with the 1RM test, which is considered the gold standard for strength measurements, showed a similar reproducibility. Horta-Gim et al. reported the absolute SEM and MDC of the 1RM test for a seated bench press in untrained 9–14-year-old participants (2.91 kg and 8.07 kg, respectively)[16]. This indicates that the 10RM test can be used as an alternative to the 1RM test for strength measurement. Conversely, Barbalho et al. reported a very small MDC in untrained elderly females (0.3 kg for both leg and bench press)[17]. We adjusted the weight in 5-kg increments, but they adjusted the weight in 500-g increments. The National Strength and Conditioning Association's 1RM test procedure involves increasing the load by 5–10 kg if the attempt is successful and decreasing by 2.5–5 kg if the attempt is unsuccessful. Although it is possible that even a higher reproducibility could have been achieved if the weight had been adjusted more precisely, we consider that adjusting the weight in 5-kg increments is sufficient.

The test-retest reliability could have been improved by adjusting the number of familiarization sessions. In this study, we conducted only one familiarization session due to concerns regarding increasing muscle strength through practice. Ritti-Dias et al. reported that for adequate evaluation of the muscular strength of young adult men with previous experience with weightlifting, at least two to three familiarization sessions were required for the 1RM test[18]. Generally, beginners take more time in familiarizing themselves with exercises than do those with training experience. An appropriate number of familiarization sessions may improve reliability. However, as the number of familiarization sessions increase, it takes more days to start the measurement, because it is necessary to leave at least 48 hours between each familiarization sessions to avoid the effects of fatigue. The degree of reliability required for muscle strength measurements remains a debate.

Regarding the results of the analysis by gender, for the females, the SEM and MDC of the 10RM test were approximately twice as large as those for males. A study by Rebeiro et al. demonstrated a greater strength change and a higher coefficient of variation in the female's 1RM test[13]. The study results are consistent with those of previous studies. However, Seo et al. reported that the 1RM testing protocol with a short warm-up and familiarization period is a reliable measurement technique for assessing muscle strength changes, regardless of the muscle group location or gender[19]. The effect of gender on muscle strength measurement remains unclear. Roberts et al. found gender differences in adaptation to strength training[20]. They stated that men may be more familiar with movements such as bench presses, even if they are beginners. Strength measurements require maximum effort, likely to strengthen training. We consider that the gender differences in potential adaptation for measurement movement affected the gender difference in measurement reliability. Further research is required to clarify this point.

The study limitations are as follows: (a) all participants were working adults, and strictly standardizing work conditions and measurement times on measurement days was not possible; (b) participants were not blinded, and motivation at the time of measurement may have influenced the measurement values; and (c) in this study, the weight of the equipment used was adjusted in 5-kg increments, which may have underestimated the participants' maximum strength.

In conclusion, this study suggests that the 10RM test could be used as an alternative to the 1RM test; however, reproducibility could be improved by blinding, finely adjusting the weight of the equipment, and unifying the conditions of the participants at the time of measurement. When conducting a 10RM test on females, understanding the possibility that the error may be greater than that in males is important.

Conflicts of interest

None.

REFERENCES

1) Alajlouni D, Bliuc D, Tran T, et al.: Decline in muscle strength and performance predicts fracture risk in elderly women and men. J Clin Endocrinol Metab, 2020, 105: 3363–3373. [Medline] [CrossRef]
2) Rubenstein LZ: Falls in older people: epidemiology, risk factors and strategies for prevention. Age Ageing, 2006, 35: ii37–ii41. [Medline] [CrossRef]
3) Riviati N, Indra B: Relationship between muscle mass and muscle strength with physical performance in older adults: a systematic review. SAGE Open Med, 2023, 11: 20503121231214650. [Medline] [CrossRef]
4) Schoenfeld BJ, Grgic J, Van Every DW, et al.: Loading recommendations for muscle strength, hypertrophy, and local endurance: a re-examination of the repetition continuum. Sports (Basel), 2021, 9: 32. [Medline] [CrossRef]
5) Fisher JP, Steele J, Gentil P, et al.: A minimal dose approach to resistance training for the older adult; the prophylactic for aging. Exp Gerontol, 2017, 99: 80–86. [Medline] [CrossRef]
6) El-Kotob R, Ponzano M, Chaput JP, et al.: Resistance training and health in adults: an overview of systematic reviews. Appl Physiol Nutr Metab, 2020, 45: S165–S179. [Medline] [CrossRef]
7) Grgic J, Lazinica B, Schoenfeld BJ, et al.: Test-retest reliability of the one-repetition maximum (1RM) strength assessment: a systematic review. Sports Med Open, 2020, 6: 31. [Medline] [CrossRef]
8) Braith RW, Graves JE, Leggett SH, et al.: Effect of training on the relationship between maximal and submaximal strength. Med Sci Sports Exerc, 1993, 25: 132–138. [Medline] [CrossRef]
9) Lavallee ME, Balam T: An overview of strength training injuries: acute and chronic. Curr Sports Med Rep, 2010, 9: 307–313. [Medline] [CrossRef]
10) Materko W, Santos EL: Prediction of one repetition maximum strength (1RM) based on a submaximal strength in adult males. Isokinet Exerc Sci, 2009, 17: 189–195. [CrossRef]

11) Reynolds JM, Gordon TJ, Robergs RA: Prediction of one repetition maximum strength from multiple repetition maximum testing and anthropometry. J Strength Cond Res, 2006, 20: 584–592. [Medline]
12) Abadie BR, Wentworth MC: Prediction of one repetition maximal strength from a 5–10 repetition submaximal strength test in college-aged females. J Exerc Physiol, 2000, 3: 1–8.
13) Ribeiro AS, do Nascimento MA, Salvador E, et al.: Reliability of one-repetition maximum test in untrained young adult men and women. Isokinet Exerc Sci, 2014, 22: 175–182. [CrossRef]
14) National Strength and Conditioning Association: Exercise technique manual for resistance training, 3rd ed. Champaign: Human Kinetics, 2016, pp 54–55, 100–101.
15) Dias Fernandes J, de Araujo Ribeiro AL, Santos W, et al.: 10-RM test of upper and lower limbs using a machine and elastic bands: Is there agreement between devices? J Phys Educ (Maringá), 2022, 33: e3319.
16) Horta-Gim MA, Romero-Perez EM, Medina-Perez C, et al.: Test-retest and minimal detectable change in the assessment of muscle strength and muscle power in upper and lower extremity exercises in 9- to 14-year-old children. Appl Sci (Basel), 2021, 11: 1–14.
17) Barbalho M, Gentil P, Raiol R, et al.: High 1RM tests reproducibility and validity are not dependent on training experience, muscle group tested od strength level in older women. Sports (Basel), 2018, 6: 171. [Medline] [CrossRef]
18) Ritti-Dias RM, Cyrino ES, Salvador EP, et al.: Influence of familiarization process on muscular strength assessment in 1-RM tests. Rev Bras Med Esporte, 2005, 11: 39–42.
19) Seo DI, Kim E, Fahs CA, et al.: Reliability of the one-repetition maximum test based on muscle group and gender. J Sports Sci Med, 2012, 11: 221–225. [Medline]
20) Roberts BM, Nuckols G, Krieger JW: Gender differences in resistance training: a systematic review and meta-analysis. J Strength Cond Res, 2020, 34: 1448–1460. [Medline] [CrossRef]