

Reliability and reproducibility of the measurement of different muscle strength manifestations

Confiabilidade e reprodutibilidade da medida de diferentes manifestações da força muscular

Daniela Bárbara Moreira Martins¹ , Ítalo Santiago Alves Viana¹ , Sandro Fernandes da Silva¹ ,
Claudia Eliza Patrocínio de Oliveira¹ , Osvaldo Costa Moreira¹ 

1. Universidade Federal de Viçosa – Campus Florestal, Florestal, MG, Brazil

ABSTRACT

Introduction: Studies that evaluate the reproducibility of CVIM and muscle power tests are scarce and often have low practical applicability. **Objective:** To evaluate the reliability and reproducibility of tests for measuring different manifestations of muscle strength. **Methods:** Nineteen healthy men, aged 23.36 ± 2.35 years, 1.82 ± 0.06 m and 80.17 ± 11.57 kg, with no experience with resistance training (RT) and no experience with the test protocols performed were evaluated. The volunteers were submitted to two assessments of maximum isometric strength (CVIM), maximum dynamic strength (1RM), and muscle power, respectively, separated by 72 hours. **Results:** For all muscle strength assessment protocols, an intraclass correlation coefficient (ICC) with high to very high strength of agreement ($ICC \geq 0.79$) was found. However, all measurements showed a moderate coefficient of variation (CV): CVIM (CV = 12.0%), mean muscle power at 40%, 60%, and 80% of the 1RM (CV = 16.2%, 11.0% and 14.0% respectively) and peak muscle power (PP) at 60% and 80% of 1RM (CV = 11.8% and 13.3% respectively), except for RM (CV = 6.4%), and PP at 40% of 1RM (CV = 5.8%), with an acceptable measurement standard error (SEM). **Conclusion:** The high to very high values for the ICC indicate an excellent reliability of the measurements in the different manifestations of muscle strength. However, since the volunteers had no experience with RT, a familiarization process prior to carrying out the tests is recommended, in order to improve their reproducibility indicators.

Keywords: health; muscle strength; muscle power.

RESUMO

Introdução: Estudos que avaliam a reprodutibilidade dos testes de CVIM e potência muscular são escassos e, muitas vezes, com baixa aplicabilidade prática. **Objetivo:** Avaliar a confiabilidade e a reprodutibilidade de testes para a medida de diferentes manifestações da força muscular. **Métodos:** Foram avaliados 19 homens saudáveis, com $23,36 \pm 2,35$ anos, com $1,82 \pm 0,06$ m e $80,17 \pm 11,57$ kg, sem experiência com treinamento de resistido (TR) e sem vivência com os protocolos dos testes realizados. Os voluntários foram submetidos a duas avaliações de força máxima isométrica (CVIM), força máxima dinâmica (1RM), e potência muscular, respectivamente, separadas por 72h. **Resultados:** Para todos os protocolos de avaliação da força muscular foi encontrado um coeficiente de correlação intraclassa (CCI) com força de concordância alta a muito alta ($CCI \geq 0,79$). No entanto, todas as medidas apresentaram coeficiente de variação (CV) moderados: CVIM (CV = 12,0%), potência muscular média a 40%, 60% e a 80% da 1RM (CV = 16,2%, 11,0% e 14,0% respectivamente) e potência muscular pico (PP) a 60% e a 80% da 1RM (CV = 11,8% e 13,3% respectivamente), à exceção da RM (CV = 6,4%), e da PP a 40% de 1RM (CV = 5,8%), com um erro padrão de medida (EPM) aceitável. **Conclusão:** Os valores altos a muito altos para o CCI indicam uma excelente confiabilidade das medidas nas diferentes manifestações de força muscular, no entanto, visto que os voluntários não tinham experiência com TR, recomenda-se um processo de familiarização prévio à realização dos testes, no sentido de melhorar os indicadores de reprodutibilidade dos mesmos.

Palavras-chave: saúde; força muscular; potência muscular.

Introduction

Strength training has been recommended for the improvement of musculoskeletal fitness, its use involves the application of high loads to provide greater adaptations such as increased strength and muscle mass [1-3]. In summary, we mention one of the principles of Strength Training, the principle of overload. In this perspective, monitoring the evolution of loads used in a strength-training program is necessary to identify changes induced by this modality.

Some strength tests have been used to monitor training loads. The maximal repetition test (1RM) is one of the most used methods for evaluating the maximum load in isoinertial movements in scientific research. According to the literature, the 1RM test is considered the gold standard in the evaluation of load-displacement through dynamic force, since it is a practical method, has a low operational cost, and with a large margin of safety for its applicability [4,5] provided that the test protocol is correctly conducted. However, its use requires some methodological care, among which pre-test familiarization has been one of the most studied. It is also evidenced in the literature that the lack of familiarization with the 1-RM test procedures may compromise the results obtained [6].

In addition, muscle strength can manifest itself in different ways: maximum, power and resistance, and these components can be expressed dynamically or isometrically [7,8]. To evaluate these different manifestations of strength, different tests can be used, such as 1RM tests, maximum voluntary isometric contraction (MVIC), and isokinetic evaluations, among others [9].

However, the literature focuses on studies on the reproducibility of the 1RM test [10-12], and studies that evaluated the reproducibility of MVIC or muscle power tests are scarce and often with low practical applicability. Thus, the evaluation of the reproducibility of protocols that evaluate different manifestations of muscle strength can add information to the scientific literature, especially on neglected tests, such as MVIC and muscle power, as well as provide practical information for the evaluation and monitoring of the different manifestations of muscle strength for professionals involved in the prescription of physical training/rehabilitation programs.

Because of lack of the literature cited above, the present study aims to evaluate the reliability and reproducibility of tests to measure different manifestations of muscle strength, more specifically, MVIC, 1RM and muscle power.

Methods

This work deals with the development of an experimental and prospective study, which comprised two phases and was carried out, in its entirety, in the Morphophysiology Laboratory of the Physical Education course of the Federal University of Viçosa Campus UFV-Florestal.

Before any experiment, the present work was approved by the Ethics Committee for Research in Human Beings of UFV (CAAE: 93793118.1.0000.5153; Opinion number: 2,919,591). In addition, all the procedures used here are by the Ethical Stan-

dards for Research in Exercise and Sports Sciences and carried out by the Declaration of Helsinki [13].

Participants

The study population consisted of university students aged between 19 and 28 years, male, and living in the city of Florestal/MG.

The inclusion criteria adopted were: being between 18 and 28 years old; being clinically able to perform the tests; not presenting any acute or chronic disease that may be affected by the performance of the tests; and consenting freely and voluntarily to perform all the procedures of the study.

Exclusion criteria were: to present any bone or joint limitation that prevents the tests; and be a user of hormonal or anti-inflammatory drugs, which may affect the outcome of the evaluations.

Volunteers who met the inclusion criteria and did not present any exclusion criteria were admitted to the study.

Considering a minimum effect size of 0.68 for muscle strength (15), a probability of α error of 0.05, and a power of (1 - β error) of 0.95, the total study sample should have at least 16 people, according to the G*Power program of the University of Düsseldorf.

Protocols and procedures

To verify the reproducibility of the measurement tests of the different manifestations of muscle strength, the MVIC, 1RM, and muscle power tests were performed.

In the evaluation of the MVIC of the lower limbs, a load cell or extensometric cell (MK, model CSL/ZL-1T, MK Controle, Brazil) with a sampling frequency of 1000 Hz was used. Before the test was carried out, the device was adjusted so that the knees of the evaluated patients were at a 90° bending angle, measured with a goniometer (Carcí, São Paulo, Brazil). At the evaluator's command, the evaluated performed a maximum isometric tension of the femoral quadriceps for 5 seconds, without letting the gluteal region lose the count with the seat so that there was no change in angulation and mechanical advantage in the lever created between the resistant force, powerful force and support point. During the execution, verbal stimuli were given to induce a higher tension, as well as the permanence of their maximum levels throughout the test. Two attempts were made, separated by an interval of 2 minutes between them, considered the highest value obtained in the two attempts [14,15].

To perform the 1RM test, the knee extension exercise was used in a BH fitness Nevada Pro-t extensor machine. The initial position adopted was similar to the MVIC test, with the individual seated with his back resting on the back of the device, his hands holding the lateral support, and the knees in 90° flexion. To perform the test, the volunteer was asked to extend the knee until it formed an angle of approximately 180° (final position) and to return to the initial position. Before the determination of 1RM, the evaluated performed the previous warm-up, which consisted of four repeti-

tions with a load of 50% of the maximum voluntary contraction value. At the end of the warm-up, the volunteer evaluated his perception of effort, using the OMNI-RES scale from 0 to 10 [16]. The load was increased at the discretion of the evaluator, according to the ease of execution and the perception of effort of the evaluated, the volunteer was asked to perform two repetitions with the new load. The load was increased until the evaluated could perform only one repetition. A maximum of five attempts could be made to determine the 1RM, with a rest interval of 2 minutes between each attempt [14,15].

The evaluation of lower limbs power was achieved through the same knee extension machine used in the MVIC and 1RM tests, starting from the same initial position (90° of kneeling flexion) and reaching the same final position (180° knee extension) of the 1RM test. Three different loads were used to evaluate the power, obtained from percentage values of 1RM (40%, 60%, and 80% of 1RM), in which the evaluated performed the knee extension movement (concentric phase of movement) at the highest possible speed. The return of the knees to the initial position occurred in a controlled manner, with a micropause of 1 to 2 seconds, to prevent the effect of accumulated elastic force from interfering in the next execution. The loads of this test were randomized in each volunteer to control a possible bias related to the learning effect or the cumulative action of fatigue. In each load, three repetitions were performed with a 2-minute rest interval between loads [15]. A linear position transducer or Chronojump Encoder (Chronojump BoscoSystem, Barcelona, Spain) was used, with a sampling frequency of 1000 Hertz, and the Chronojump Software, version 1.6.2, (Chronojump BoscoSystem, Barcelona, Spain), to determine the power values. Through this instrument, it was possible to obtain information about average power (AP) and peak power (PP).

The experimental design adopted in this study allowed the evaluation of the reliability of the reproducibility of the tests to measure the different manifestations of muscle strength. For this, all the evaluated patients underwent two evaluations, separated by 72h, in which they performed the same procedures. From these evaluations, data regarding 1RM, MVIC, AP, and PP of each participant were analyzed. These outcome parameters were used to calculate the coefficient of variation (CV), the intraclass correlation coefficient (ICC), and its 95% confidence interval (CI95%).

Statistical treatment

All statistical analyses were performed in the statistical program SPSS for Windows, version 23 (IBM, Chicago, USA). Initially, the data were submitted to the Shapiro-Wilk test to verify normality and, later, generated means and standard deviations (SD) for descriptive analysis of the data. The reliability of the measurement tests of the different manifestations of muscle strength will be determined by the calculation of CV and ICC. The ICC was also used to verify the reproducibility of muscle strength measurements, with confidence interval of 95% (CI95%). ICC values equal to or greater than 0.90 can be considered very high, values between 0.70 and 0.89, can be

considered high and values between 0.50 and 0.69, are moderate [17,18]. The values for the standard error of the measurement (SEM) were also calculated by the product of the basal standard deviation with the square root of (1-r), in which r is the intra-class correlation ($SEM = SD \times \sqrt{1-ICC}$). In addition, the minimally detectable change (MDC) by the SEM product was calculated with the square root of two (due to the variance of the measurement error of each instrument) and with the value of 1.96, which represents that of a normal standard curve associated with a 95% confidence interval ($MDC = 1.96 \times \sqrt{2} \times SEM$) [18,19]. A statistical significance level of $p < 0.05$ was established for all treatments.

Results

The present study participated in 19 male volunteers, with no previous experience with strength training or who did not train for at least six months, with a mean age of 23.36 ± 2.35 years, mean height of 1.82 ± 0.06 m, and mean body mass of 80.17 ± 11.57 kg. During the tests, there were no reports of osteomyoarticular or metabolic problems related to their performance.

Data on the reliability and reproducibility of test protocols to measure the different manifestations of muscle strength can be observed in Table I. It is possible to observe, from the values found, that for all protocols for evaluating the different manifestations of muscle strength, a high to very high ICC was found (≥ 0.79). However, all measures presented moderate CV (except for 1RM and PP at 40% of 1RM), with an acceptable SEM.

Table I - Results for reliability and reproducibility of measurements obtained in test protocols for different manifestations of muscle strength

	Mean	SD	Mean	SD	CV	ICC (CI 95%)	p	SEM	MDC
MVIC (kg)	104.58	23.61	98.37	22.25	12.0%	0.94 (0.83;0.98)	<0.001	5.45	15.11
1RM (kg)	86.87	14.58	84.75	15.02	6.4%	0.97 (0.91;0.99)	<0.001	2.60	7.21
AP40 (w)	1092.36	199.13	1102.95	198.31	16.2%	0.79 (0.21;0.94)	0.011	90.88	251.90
PP40 (w)	2614.07	469.13	2544.55	410.39	5.8%	0.97 (0.91;0.99)	<0.001	71.08	197.03
AP60 (w)	951.33	158.91	938.51	159.99	11.0%	0.89 (0.66;0.97)	<0.001	53.06	147.08
PP60 (w)	1985.33	414.86	1868.27	362.55	11.8%	0.91 (0.73;0.97)	<0.001	108.77	301.48
AP80 (w)	682.30	155.31	675.67	168.83	14.0%	0.92 (0.74;0.98)	<0.001	47.75	132.36
PP80 (w)	1328.35	323.28	1331.84	359.21	13.3%	0.93 (0.77;0.98)	<0.001	95.04	263.43

SD = standard deviation; CV = coefficient of variation; ICC = intraclass correlation coefficient; 95% CI = 95% confidence interval; SEM = measurement standard error; MDC= minimally detectable change; MVIC = maximal isometric strength; 1RM = maximal dynamic strength; AP40 = muscle power average at 40% of 1RM; PP40 = peak of muscle power at 40% of 1RM; AP60 = muscle power average at 60% of 1RM; PP60 = peak of muscle power at 60% of 1RM; AP80 = muscle power average at 80% of 1RM; PP80 = peak of muscle power at 80% of 1RM

Although the CV value found was moderate for the results of MVIC (12%), the ICC found for the MVIC test (0.94) is in line with the results obtained by Neves *et al.* [20] who evaluated 30 volunteer students (of both sexes and age between 18 and 25 years) in the knee extension exercise in the flexo-extensor chair. For the collection of the results, a load cell of 250 kg positioned in the flexor chair and connected to the Software Miotool USB 400 was used, which translated the values to the computer. At the applicator command, the evaluated person was instructed to perform maximum voluntary isometric force and verbally stimulated to maintain the maximum levels for 6 seconds. At the end of the time he was instructed to relax and rest for 1 minute, until the next attempt, six attempts were made for each lower limb, three attempts commanded by an evaluator and three commanded by another evaluator. The protocol was repeated after 48h. The analysis of the values was performed by Cronbach's alpha and the results showed a strong correlation on the first day ($\alpha= 0.980$) on the second day they found values similar to the first ($\alpha= 0.982$), this means that the test had an excellent reliability index.

Although there are differences between the protocol used by Neves *et al.* [20] and the protocol we used in this study, both provided relevant data for the literature, with a high degree of reliability and reproducibility. We emphasize that for better and more reliable results it is recommended to familiarize with the test and with the movement that will be performed, according to Brown and Weir [9] and more the strength test has generally shown to be reliable, novice individuals are likely to improve their scores in subsequent tests simply due to familiarization and comfort during the test.

Verdik *et al.* [21] prove the reproducibility of the 1RM test as a valid measure of knee extensor strength, regardless of physical conditioning and age of the subject. Although the 1RM test is low operational cost and easy to reproducibility, according to Dias *et al.* [22] some points should be taken into consideration before its execution, such as: starting the test rested and effortlessly strenuous for at least 24 hours prior to the test; have knowledge of the movement that will be performed and if you have not had an experience that goes through a learning session; be motivated to achieve your best performance and consequently its best result; perform previous heating; not let 1RM attempts pass 5; and that there is a recovery interval between attempts so that there is no energy impairment in subsequent ones. Given the low CV value found in the 1RM test results (6.4%), it is possible to consider that, from the results found, it has good applicability, reliability, and reproducibility.

The CV and ICC values found for PP at 40% appear to be similar to those found by Pagaduan and Blas [23] who tested 15 college students from the following protocol: perform a jump as high as possible against a contact platform and with a 20 kg bar placed on the shoulders, in order to verify the reliability of a movement jump loaded using the Chronojump-BoscoSystem and thus establish the reproducibility of the test to aggregate the literature related to lower-lower-heel potency. The results were obtained from the Chronojump-BoscoSystem Software and the statistical

analysis they found the ICC considered moderate to high (0.86) and a low value for CV (6.7%), which fits into a reliable test according to the software used. We emphasize that studies evaluating the reproducibility of power tests are scarce and make it difficult to directly compare the results obtained in our study with those obtained in other studies. Although it is not the same protocol and the same test used in our study, we evidenced the high potential that these tests have to aggregate the literature to measure the different types of strength already mentioned.

The values found for CV in the 1RM test and the PP test at 40% of 1RM are considered low, 6.4%, and 5.8%, respectively, presenting excellent reproducibility. On the other hand, the other protocols presented values between 11% and 16%, and moderate reproducibility was considered. According to Atkinson and Nevill [24], absolute reproducibility has been considered acceptable when the value found is below 10%. However, it is worth noting that this change in the coefficient of variation can be explained by some points such as lack of familiarization with the test, fatigue, and the load used in the two days of testing.

Determining SEM measures is crucial in order to be able to compare the measures between the evaluations. To measure and evaluate the reliability of both protocols we used SEM and MDC in all values, and from the results, we can say that the tests performed presented excellent reliability measures and good results for the reproducibility of the tests of the different manifestations of muscle strength.

Conclusion

Because of the high to very high values for the ICC, it can be concluded that the present study demonstrates that, even in beginners, muscle strength assessment tests present excellent reliability of measurements in the different manifestations of muscle strength. However, a CV was found with considerable variability, indicating that, because they are volunteers without previous experience with RT, a familiarization process prior to the tests is recommended, in order to improve their reproducibility indicators.

Conflict of interest

The authors declare that there are no conflicts of interest.

Financing

Funded by the FAPEMIG (Research Support Foundation of Minas Gerais State).

Authors' contributions

Conception and design of the research: Martins DBM, Moreira OC; **Obtaining data:** Martins DBM, Moreira OC, Santiago IA; **Analysis and interpretation of data:** Martins DBM, Moreira OC; **Statistical analysis:** Martins DBM, Moreira OC; **Writing of the manuscript:** Moreira OC, Martins DBM, Santiago AI, Oliveira CEP; **Critical revision of the manuscript regarding intellectual content important:** Oliveira CEP, Silva SF.

References

1. American College of Sports Medicine. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sports Exerc.* 2009;41(2):459-71. doi: 10.1249/mss.0b013e3181949333
2. American College of Sports Medicine. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc.* 2011;43(7):1334-59. doi: 10.1249/mss.0b013e318213fefb
3. Santos LV, Pereira ET, Reguera-García MM, Oliveira CEP, Moreira OC. Resistance training and muscle strength in people with spinal cord injury: a systematic review and meta-analysis. *J Bodyw Mov Ther.* 2022;29:154-60. doi: 10.1016/j.jbmt.2021.09.031
4. Rhea MR, Alvar BA, Burkett LN, Ball SD. A meta-analysis to determine the dose response for strength development. *Med Sci Sports Exerc.* 2003;35(3):456-64. doi: 10.1249/01.MSS.0000053727.63505.D4
5. Mayhew JL, Johnson BD, Lamonte MJ, Lauber D, Kemmler W. Accuracy of prediction equations for determining one repetition maximum bench press in women before and after resistance training. *J Strength Cond Res.* 2008;22(5):1570-7. doi: 10.1519/JSC.0b013e31817b02ad
6. Rydwik E, Karlsson C, Frändin K, Akner G. Muscle strength testing with one repetition maximum in the arm/shoulder for people aged 75 + - test-retest reliability. *Clin Rehabil.* 2007;21(3):258-65. doi: 10.1177/0269215506072088
7. Knuttgen HG. Measurement and terminology for authors and reviewers. *J Sports Med Phys Fitness.* 2019;59(2):341-2. doi: 10.23736/S0022-4707.18.09363-5
8. Moreira OC, Oliveira CEP, Maroto-Izquierdo S, Cuevas MJ, Paz JA. Effects of short-term strength training on body composition, muscle strength and functional capacity of elderly: a systematic review and meta-analysis. *Biosci J.* 2019;35(6):1941-57. doi: 10.14393/BJ-v35n6a2019-42775
9. Brown LE, Weir JP. Asep procedures recommendation I: accurate assesment of muscular strength and power. *JEPonline.* 2001;4(3):1-21.
10. Dias RMT, Cucato GG, Camara LC, Wolosker N. Reprodutibilidade do teste de 1-RM em indivíduos com doença arterial obstrutiva periférica. *Rev Bras Med Esporte.* 2010;16(3):201-4. doi: 10.1590/s1517-86922010000300009
11. Mitter B, Csapo R, Bauer P, Tschan H. Reproducibility of strength performance and strength-endurance profiles: A test-retest study. *PLoS One.* 2022;17(5):e0268074. doi: 10.1371/journal.pone.0268074
12. Schoenell MCW, Tiggemann CL, Cadore EL, Tartaruga MP, Kruegel LFM. Correlação e reprodutibilidade de testes abdominais em mulheres jovens. *Rev Bras Ciênc Esporte.* 2013;35(3):561-574. doi: 10.1590/s0101-32892013000300003
13. Williams JR. The Declaration of Helsinki and public health. *Bull World Health Organ.* 2008;86(8):650-2. doi: 10.2471/blt.08.050955
14. Oliveira CEP, Moreira OC, Carrión-Yagual ZM, Medina-Pérez C, Paz JA. Effects of classic progressive resistance training versus eccentric-enhanced resistance training in people with multiple sclerosis. *Arch Phys Med Rehabil.* 2018;99(5):819-25. doi: 10.1016/j.apmr.2017.10.021
15. Moreira OC, Cardozo RMB, Vicente MA, Matos DG, Mazini Filho ML, Guimarães MP, et al. Acute effect of stretching prior to resistance training on morphological, functional and activation indicators of skeletal muscle in young men. *Sport Sci Health.* 2022;18:193-202. doi: 10.1007/s11332-021-00793-0
16. Robertson RJ, Goss FL, Rutkowski J, Lenz B, Dixon C, Timmer J, et al. Concurrent validation of the OMNI perceived exertion scale for resistance exercise. *Med Sci Sports Exerc.* 2003;35(2):333-41. doi: 10.1249/01.MSS.0000048831.15016.2A
17. Bosquet L, Maquet D, Forthomme B, Nowak N, Lehance C, Croisier JL. Effect of the lengthening of the protocol on the reliability of muscle fatigue indicators. *Int J Sports Med.* 2010;31(2):82-8. Epub 2009 Dec 17. doi: 10.1055/s-0029-1243168
18. Mentiplay BF, Perraton LG, Bower KJ, Adair B, Pua YH, Williams GP, et al. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: a reliability and validity study. *PLoS One.* 2015;10(10):e0140822. doi: 10.1371/journal.pone.0140822
19. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res.* 2001;10(7):571-8. doi: 10.1023/a:1013138911638
20. Neves CDC, Tossige-Gomes R, Avelar NCP, Simão AP, Lacerda ACR. Avaliação da confiabilidade da força isométrica de extensores de joelho pelo uso da célula de carga. *Ter Man.* 2011;9(41):16-21.
21. Verdijk LB, van Loon L, Meijer K, Savelberg HH. One-repetition maximum strength test represents a valid means to assess leg strength in vivo in humans. *J Sports Sci.* 2009;27(1):59-68. doi:

10.1080/02640410802428089

22. Dias RMR, Avelar A, Meneses LA, Salvador PEM, Da Silva PRD, Cyrino ED. Segurança, reprodutibilidade, fatores intervenientes e aplicabilidade de testes de 1-RM. *Motriz*. 2013;19(1):231-42. doi: 10.1590/S1980-65742013000100024

23. Pagaduan CJ, Blas X. Reliability of a loaded countermovement jump performance using the chronojump-boscosystem. *Kinesiologia Slovenica*. 2012;18(2):45-8.

24. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998;26(4):217-38. doi: 10.2165/00007256-199826040-00002

