Development and assessment of test-retest reliability of a new field test to evaluate lower-limb muscle fatigability in young adults

Robin Souron[1,2,*], Julian Colard[1,2], Juan D. Ruiz-Cárdenas[3], Anicée Beltran[1], Pascale Duché[1], and Mathieu Gruet[1]

[1] Université de Toulon, J-AP2S, Toulon, France
[2] Nantes Université, Movement Interactions Performance, MIP, UR 4334, F-44000 Nantes, France
[3] Physiotherapy Department, Faculty of Physiotherapy, Podiatry and Occupational Therapy, Universidad Católica de Murcia, Murcia, Spain

Abstract – We aimed to develop a new field test to evaluate lower-limb muscle fatigability in young adults. In Experiment–A, we developed and determined the ability of an intermittent isometric wall-squat test to induce progressive level of muscle fatigability, as detected by the kinetics of changes in squat-jump height (SJ_H) and sit-to-stand time (STS_T) computed using two smartphone applications for feasibility purposes. In Experiment–B, participants performed the same test on two different days for reliability assessment. Kinetics of changes in our fatigability indicators were registered at isotime, exhaustion, and $Post_{2min}$. The minimal detectable change (MDC_{95}) and the absolute (CV_{TE}) and relative ($ICC_{3\text{-}1}$) reliability coefficients were assessed. In Experiment–A, we reported a progressive decrease in performance for SJ_H and STS_T throughout the task, reaching at exhaustion mean changes of $-22\pm11\%$ and $+31\pm13\%$. Individual data-analysis showed decrease in performance for SJ_H and STS_T greater than the MDC_{95} in 85% and 95% of participants. In Experiment–B, changes in our fatigability indicators demonstrated excellent inter-session reliability at isotime, exhaustion and $Post_{2min}$ for SJ_H ($ICC_{3\text{-}1} > 0.97$; $CV_{TE} < 7.5\%$) and STS_T ($ICC_{3\text{-}1} > 0.92$; $CV_{TE} < 3.3\%$). This test is feasible and reliable, making it very promising for evaluating muscle fatigability in applied (e.g. clinical) and laboratory settings.

Keywords: Performance fatigability, muscle testing, muscle power, muscle endurance, smartphone applications

Résumé – Développement et évaluation de la reproductibilité test-retest d'un nouveau test de terrain visant à évaluer la fatigabilité des muscles des membres inférieurs chez les jeunes adultes. L'objectif de cette étude était de développer un nouveau test écologique de terrain pour évaluer la fatigabilité des muscles des membres inférieurs chez des jeunes adultes. Dans l'expérience A, nous avons conçu et évalué la capacité d'un test intermittent impliquant l'exercice de chaise isométrique pour induire un niveau progressif de fatigabilité musculaire, détectée par la cinétique des changements de hauteur de saut en squat jump (SJ_H) et de temps de lever de chaise (STS_T), calculée à l'aide de deux applications mobiles à des fins de faisabilité. Dans l'expérience B, les participants ont effectué le même test lors de deux journées différentes pour évaluer la reproductibilité test-retest. Les cinétiques des changements de nos indicateurs de fatigabilité ont été enregistrées à *isotime*, à l'épuisement et 2 min après l'arrêt de l'exercice. La variation minimale détectable (MDC_{95}) ainsi que les coefficients de fiabilité absolue (CV_{TE}) et relative ($ICC_{3\text{-}1}$) ont été évalués. Dans l'expérience A, nous avons observé une diminution progressive des performances pour le SJ_H et le STS_T tout au long de la tâche, atteignant à l'épuisement des changements moyens de $-22\pm11\%$ et $+31\pm13\%$. L'analyse des données individuelles a montré une diminution des performances pour le SJ_H et le STS_T supérieure au MDC95 chez 85% et 95% des participants. Dans l'expérience B, les changements dans nos indicateurs de fatigabilité ont montré une excellente fiabilité inter-session à *isotime*, à l'épuisement et 2 min après l'arrêt de l'exercice pour le SJ_H ($ICC_{3\text{-}1} > 0{,}97$; $CV_{TE} < 7{,}5\%$) et le STS_T ($ICC_{3\text{-}1} > 0{,}92$; $CV_{TE} < 3{,}3\%$). Ce test est réalisable et fiable, ce qui le rend particulièrement prometteur pour évaluer la fatigue musculaire, que ce soit dans des environnements appliqués tels que la clinique ou en laboratoire.

Mots clés : Fatigabilité, testing musculaire, puissance musculaire, endurance musculaire, applications mobiles

1 Introduction

Muscle endurance is an important component of neuromuscular function and refers to the ability of a muscle or a muscle group to perform repeated work (Caspersen, Powell, & Christenson, 1985). The main outcomes informing on muscle endurance are either the total amount of work performed or the time to exhaustion in the case where the task requires to maintain or repeat a given level of voluntary contraction for as long as possible (Evans *et al.*, 2015). While endurance time is a relevant indicator of muscle performance for such tests, it also greatly depends on psychological factors such as motivation and the ability to withstand pain, which exert an increasing influence as participants gets close to exhaustion. Therefore, it is useful to supplement this outcome with other indicators of muscle performance collected before symptom limitation. Within this framework, it is possible to incorporate performance fatigability indicators early in the test and at regular interval until exhaustion. Performance fatigability refers to the kinetics of change in a given performance indicator, e.g., muscle strength or power (Enoka & Duchateau, 2016). For instance, many studies used changes in maximal muscle strength of a muscle group to track the kinetics of muscle fatigability development in various contexts and populations (Bachasson, Millet *et al.*, 2013; Gruet *et al.*, 2016; Souron et al., 2020). These tests generally assess maximal muscle strength at regular intervals throughout the exercise protocol, offering insights on muscle functioning across different effort levels and fatigue states, which may be pertinent to various activities of daily living. Overall, endurance time, fatigability assessments throughout the test and fatigability level at volitional exhaustion are complementary to provide a comprehensive assessment of neuromuscular performance during prolonged motor tasks. Bachasson, Millet *et al.* (2013) developed a hybrid protocol allowing to compute all these outcomes from a prolonged motor task consisting in intermittent isometric knee extensions performed at increasing force levels until exhaustion, interspaced with muscle fatigability assessments at regular intervals. Despite its numerous advantages, this test shares the common drawbacks of other fatiguing tests requiring specific ergometers, load cells/strain gauges and dedicated software, with large space requirements and elevated costs usually restricting their use to laboratory settings.

Here we sought to develop an innovative test allowing the assessment of both lower-limb muscle fatigability and endurance and gathering several advantages of more traditional fatigability protocol while remaining highly feasible and affordable. We considered the use of the isometric wall squat as the fatiguing exercise. This choice was guided by the fact that such exercise procedure i) is easy to standardize and replicate (Lea, O'Driscoll, Coleman, & Wiles, 2021), ii) does not require specific material as opposed to other tests (Bachasson, Millet *et al.*, 2013), iii) targets various lower-limb muscle groups, i.e., quadriceps, hamstrings, gluteal (Ayotte, Stetts, Keenan, & Greenway, 2007), iv) isolates the neuromuscular function with low cardiorespiratory requirements (Lea *et al.*, 2021), v) can be implemented with a progressive difficulty by adjusting the duration of the exercise, vi) can be conducted in an intermittent fashion and interrupted without delay, allowing to incorporate regular fatigability assessments and vii) is feasible in various populations, including the general population, athletes and several clinical populations (Baffour-Awuah *et al.*, 2023 Baffour-Awuah, Pearson, Dieberg, Wiles, & Smart, 2023; Goldring, Wiles, & Coleman, 2014; Wiles *et al.*, 2018 Wiles, Taylor, Coleman, Sharma, & O'Driscoll, 2018). We finally considered the use of simple indicators of muscle performance based on tests relevant to daily life (i.e., vertical jump and sit-to-stand) to track changes in muscle fatigability.

This work aimed to develop a new field test to evaluate lower-limb muscle fatigability in young adults, as detected by the kinetics of changes in squat-jump height (SJ_H) and sit-to-stand time (STS_T) computed using two smartphone applications (experiment A), and to assess its test-retest reliability in an independent sample of young adults (experiment B). We hypothesized that the test we developed in this work would induce a progressive development of muscle fatigability and that changes in the indicators computed to evaluate muscle fatigability would demonstrate a high test-retest reliability.

2. Materials and methods (experiment A)

2.1 Study design

Two independent experiments were built in a logical fashion. The first one (experiment A) aimed to determine if muscle fatigability could be detected from the kinetic of change in SJ_H and STS_T along the course of an incremental isometric exercise performed until exhaustion. The second one (experiment B) aimed to evaluate the test-retest reliability of the muscle fatigability test on an independent sample of participants. We planned to adapt the protocol for experiment B (e.g., duration of the stages, deletion of one measure if no change was detected for its associated outcome during the protocol) based on the results obtained during experiment A. For both experiments, none of the participants had known cardiovascular, musculoskeletal, auditory, or cognitive disorders that may impair their ability to perform the test. Participants were asked to refrain from strenuous and unaccustomed physical activity at least 24h prior testing. Participants were informed of the benefits and risks of the investigation prior to signing an institutionally approved informed consent to participate in the study. This study conformed to standards from the latest revision of the Declaration of Helsinki and was approved by a national research ethic committee (n° IRB00012476-2021-08-02-87).

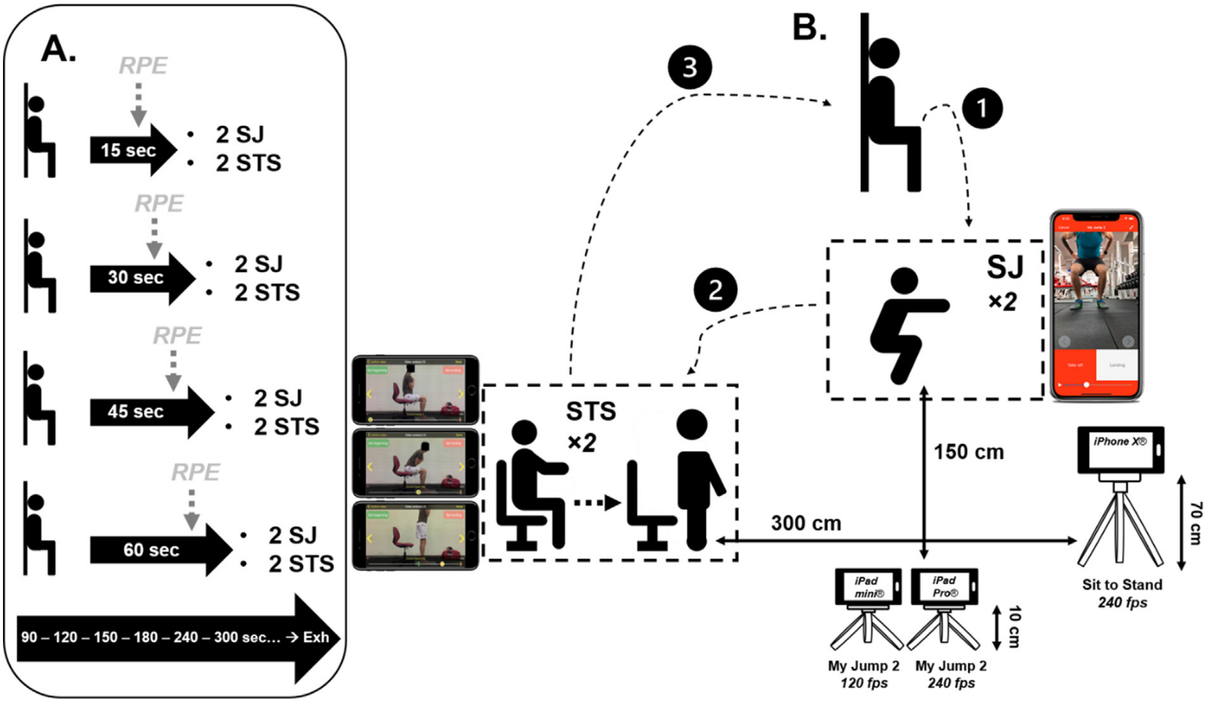

Fig. 1. Panel A schematically represents the muscle fatigability protocol. The duration of the fatiguing exercise was progressively increased until exhaustion (Exh), that was defined as the moment when participants voluntarily stopped the exercise or the moment when participants were not able to maintain a correct wall squat position for more than three consecutive seconds. At the end of each stage, participants performed two squat jumps (SJ) and two sit-to-stand (STS) exercises. Panel B displays the experimental layout. The fatigability protocol began with the first 15-s stage. Ten second before the end of each stage, participants were asked to verbally report their perception of muscular and respiratory efforts (dotted grey arrow). At the end of each exercise stages, participants moved from the wall squat position to the SJ area (1). After performing two squat jumps, participants moved to the STS area and performed two STS exercises (2). Then, participants returned to the wall squat position to start a new fatiguing stage (3), and so on until exhaustion. Screenshot images for Sit to Stand and My Jump 2 applications were re-used with permission from Ruiz-Cárdenas and colleagues (Ruiz-Cardenas *et al.*, 2018) and from http://carlos-balsalobre.com, respectively).

2.2 Participants

We recruited twenty participants (eight females; mean ± SD age: 24 ± 2 years, height: 173 ± 8 cm, weight: 69 ± 7 kg) who visited the laboratory on two occasions, i.e., one familiarization and one experimental session. All participants were recreationally active, which reflected a commitment of 2–4 h of physical activity per week, i.e. level 2 of performance according to guidelines from De Pauw *et al.* (2013). During the familiarization session, participants were accustomed with all the procedures of the experimental session, including the procedures for the good execution of sit-to-stand and squat jump exercises, the fatigability protocol (i.e., completion of the first three stages of isometric wall squat, see below) and the report of rating of perceived exertion (RPE) scores.

2.3 Procedures

2.3.1 Squat jump performance

In line with our reasoning of developing a highly feasible test for optimal implementation, SJ_H was evaluated with a valid and reliable mobile App called *My Jump 2* installed on an iPad mini® (Apple Inc., California, USA) with a 120-fps recording mode (Balsalobre-Fernandez *et al.*, 2015 Balsalobre-Fernandez, Glaister, & Lockey, 2015). Participants started the jump with the feet a shoulder-width apart, the knee joint at approximately 90° and with their hands placed on their hips (Balsalobre-Fernandez *et al.*, 2015). The iPad was positioned in front of the participants, at 150 cm of his feet, enabling a clear view of the participant's lower limbs to ascertain take-off and post jump landing and any possible non-desirable counter movement (Balsalobre-Fernandez *et al.*, 2015). SJ_H was calculated by manually selecting the take-off (i.e., the first frame in which both feet were off the ground) and the landing (i.e., the first frame in which at least one foot was touching the ground; see Figure 1B or http://carlos-balsalobre.com/#apps for a zoomed display) to estimate the flight time. Jump height was calculated using the following equation:

$SJ_H = t^2 \times 1.22625$, with SJ_H being the jump height in cm and t representing the flight-time in s.

2.3.2 Sit-to-Stand performance

We calculated STS_T using the *Sit to Stand* application (Ruiz-Cardenas *et al.*, 2018Ruiz-Cardenas, Rodriguez-Juan, Smart, Jakobi, & Jones, 2018) installed on an iPhone X (240 fps; Apple Inc., California, USA). For the sit-to-stand test, we used a standard chair (46 cm height) with a flat seat and no armrest. Participants were asked to sit with their legs hip-width apart and with hip, knee and ankle joints at ~90° (Strassmann *et al.*, 2013). Participants started in a sitting position and were asked to stand-up once as fast as possible (i.e., one transition from sitting to standing) while maintaining their arms crossed over the chest. The mobile device was positioned on a 0.7 m high tripod placed 3 m from the right side of the participants (Fig. 1B) (Ruiz-Cardenas *et al.*, 2018). This mobile application was designed to analyse sit-to-stand test via high-speed video recording (i.e., 240 fps) to allow the calculation of the rising time (in s) between two frames manually selected by the experimenter (Fig. 1B).

A 90-s resting period separated the baseline measurements (i.e., the three squat jumps and sit-to-stand) from the beginning of the fatigability protocol.

2.3.3 Test protocol

The whole protocol is illustrated in Figure 1. The selected fatigability task consisted in maintaining an isometric wall squat position. First, posture was adjusted to obtain knee and hip angles at 100° using a clinical goniometer, with participants keeping their lower legs and trunk vertical (i.e., back against the wall). A mark was then placed on the midline of the femur and a laser system (CL2i STANLEY) was used to ensure the maintenance of the position. A wooden board was placed on the floor just in front of the feet of the participants. Both the laser system and the wooden board allowed for rapid and accurate replication of the position at the beginning of each new stage. Participants were instructed to wear sport shoes. A 50 × 50 cm non-slippery surface was positioned under the feet of participants to ensure maximal adherence to the floor.

The duration of the isometric wall squat task was progressively incremented (i.e., 15-s incremental stages for the four first stages, 30-s incremental stages for the fifth to eighth stages and 60-s incremental from the ninth stages) until exhaustion (Fig. 1A). Exhaustion (Exh) was the moment where participants voluntarily stopped the exercise or the moment where they were not able to maintain the wall squat position for more than three consecutive seconds. Standardized sentences (e.g., "you're doing great, keep going") were pronounced by the experimenter with a neutral tone at the middle and 15 s before the end of each stage. In a 30-s window, participants performed two squat jumps and two sit-to-stand exercises immediately at the end of each stage, at exhaustion and 2 min after exhaustion ($Post_{2min}$).

We used *My Jump 2* (Balsalobre-Fernandez *et al.*, 2015) and *Sit to Stand* (Ruiz-Cardenas *et al.*, 2018) mobile Apps to assess SJ_H and STS_T. Internal validation showed an excellent intra-class correlation coefficient two-way mixed effect absolute agreement for the two squat jumps ($ICC_{2\text{-}k} = 0.98$) and the two sit-to-stand tests ($ICC_{2\text{-}k} = 0.93$) prior to the fatigability protocol. The kinetics of changes in these indicators were used as indices of muscle fatigability. The endurance time was calculated by summing the time of each completed stages and the time performed in the last stage, i.e., the one where the participants reached exhaustion. Participants were asked to verbally report their perception of muscular (RPE_M) and respiratory (RPE_R) efforts (Borg 6–20 scale) 10 s before the end of each stage. Similar to studies employing exercise modalities expected to predominantly activate specific systems (e.g. limb muscles during isolated isometric exercises), we opted for the use of differentiated RPE scores (Bolgar, Baker, Goss, Nagle, & Robertson, 2010; Borg, 1998; Robertson & Noble, 1997), using the 6–20 RPE scale that has been used repeatedly for that purpose [e.g. Green, Crews, Bosak and Peveler (2003); Gruet et al., 2018 Gruet, Mely and Vallier (2018); Pageaux et al., 2013 Pageaux, Marcora and Lepers (2013)]. The heart rate (HR, in bpm) was monitored throughout the fatigability test using a Polar Vantage M watch (Polar Electro, Kempele, Finland). The peak HR value (HR_{PEAK}) obtained during the fatigability test was expressed as a percentage of the theoretical maximal heart rate (HR_{MAX}). Theoretical HR_{MAX} was calculated using the following equation (Gellish *et al.*, 2007): $HR_{MAX} = 192–0.007 \times age^2$

2.4 Statistical analysis

Sample size was estimated to detect differences between repeated measures with a moderate effect size (F test ≥ 0.27), an alpha level of 0.05, and a desired statistical power of 80% for a one-tailed test. A minimum of 16 participants was estimated. Normal distribution and homogeneity of variances analysis were confirmed using the Kolmogorov–Smirnov and Skewness test, respectively.

A group isotime analysis was performed (Nicolò *et al.*, 2019). This includes only the portion of the fatiguing test that was available for all participants, which is limited by participants with the shorted exercise time (i.e., the lowest number of completed stages). The lowest number of completed stages was three (S1, S2 and S3 for the first, second and third stages, respectively). In order to analyse changes in our fatigability indicators (i.e., SJ_H, STS_T, HR, RPE_M, RPE_R) throughout the fatiguing protocol, a one-way repeated measures ANOVA with the fatigability indicators as dependent variable and time (i.e., Pre, S1, S2, S3, Exh, $Post_{2min}$) as independent variable was used. If significant main effects were observed, these were followed up by post hoc Bonferroni-corrected pairwise comparisons. Partial eta square (η_p^2) was reported as an estimate of effect size, with $\eta_p^2 \geq 0.07$ and $\eta_p^2 \geq 0.14$ used as moderate and large effects, respectively (Cohen, 1988). Additionally, the minimal detectable change with 95% confidence (MDC_{95}) was calculated as $SEM \times \sqrt{2} \times 1.96$ and expressed as absolute and percentage of change (Weir, 2005). The MDC_{95} can be interpreted as the smallest change in a measure that can be considered real change beyond measurement error with 95% confidence. The alpha level for all statistical tests was set a $p < 0.05$.

Table 1. Heart rate (HR), perceptions of muscular (RPE_M) and respiratory (RPE_R) efforts recorded during the first three stages of the test (S1, S2 and S3, respectively) and at exhaustion (Exh).

	S1	S2	S3	Exh
HR (bpm)	124 ± 15	147 ± 12***	157 ± 15***	170 ± 15***
RPE_M (6–20 scale)	8.4 ± 1.7	10.8 ± 2.0***	13.4 ± 2.5***	19.1 ± 1.7***
RPE_R (6–20 scale)	6.8 ± 0.7	8.4 ± 1.6*	9.9 ± 2.3***	13.6 ± 2.8***

Perceptions of muscular and respiratory efforts were measured using the Borg' 6–20 scale.
Significantly different from S1: * p <0.05, *** p < 0.001.

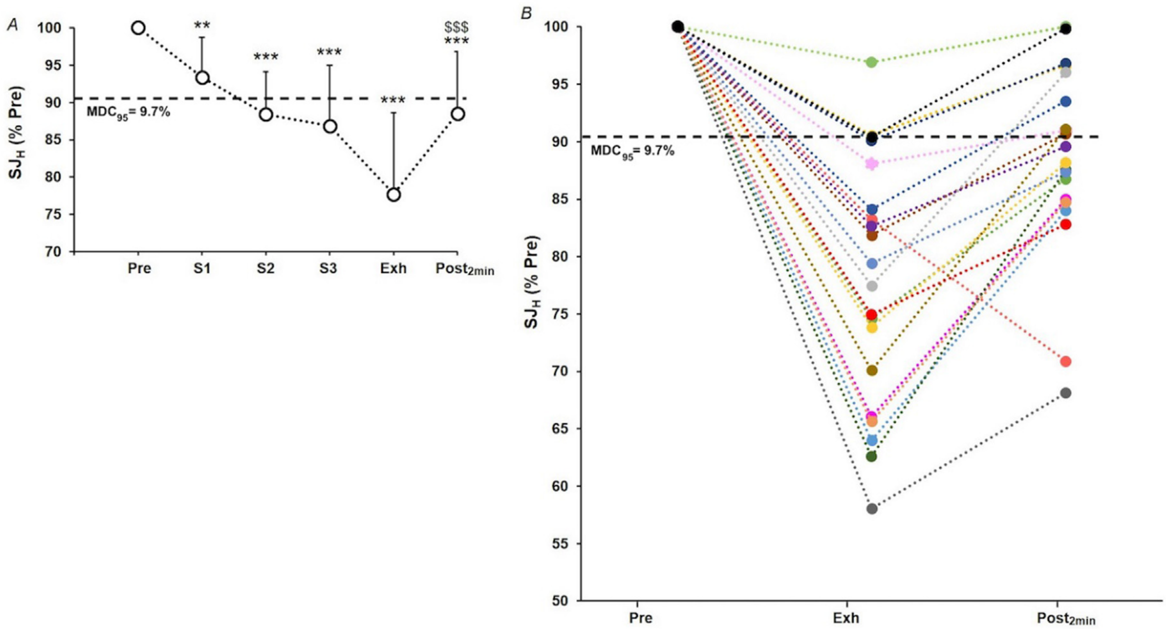

Fig. 2. Squat jump height (SJ_H) recorded before (Pre), during (S1, S2, S3), at exhaustion (Exh) and 2 min after the end of the fatiguing test ($Post_{2min}$). Panel A displays the pooled data of the 20 participants. Panel B displays the individual data for values recorded before the protocol (Pre), at exhaustion (Exh) and after a 2-min recovery period ($Post_{2min}$). Error bars on the panel A denote standard deviation. Horizontal dashed line represents the minimal detectable change (MDC_{95}). Data are expressed as percentage of Pre. Significantly different from Pre: ** $p < 0.01$, *** $p < 0.001$. Significantly different from Exh: $$$ $p < 0.001$

3 Results (experiment A)

3.1 Endurance time, heart rate and rating of perceived exertion

The mean number of completed stages was 5.3 ± 2.0 (range 3–10 stages), with a mean exercise duration of 398 ± 339 s (range 110–1429 s). A significant time effect was found for HR (p < 0.001; $p\eta^2 = 0.85$) with a progressive increase throughout the fatiguing exercise (Tab. 1). The mean HR_{PEAK} obtained during the test was 91 ± 7% of theoretical HR_{MAX} (range 78–100%). Significant time effects were found for RPE_M (p < 0.001; $\eta_p^2 = 0.93$) and RPE_R ($p < 0.001$; $\eta_p^2 = 0.78$), with values recorded at Exh significantly different from values recorded at the end of the first stage ($p < 0.001$; Tab. 1).

3.2 Squat jump performance

A significant time effect was found for SJ_H ($p < 0.001$; $\eta_p^2 = 0.62$). SJ_H progressively decreased throughout the fatiguing protocol, with mean decreases of 6.6 ± 5.4, 11.6 ± 5.7, 13.2 ± 8.2 and 22.3 ± 10.9% at S1, S2, S3 and Exh when compared to Pre, respectively (Fig. 2A). Mean of changes in SJ_H was greater than the MDC_{95} (2.7 cm; 9.71%) for S2, S3 and Exh. Two minutes of recovery were not sufficient for SJ_H to fully recover, with a 11.5 ± 8.3% of mean decrease in performance at $Post_{2min}$ being greater than the MDC_{95}. However, this time frame was sufficient to induce a partial recovery, with SJ_H measured at $Post_{2min}$ being significantly different from Exh ($p < 0.001$) and non-different from S1 ($p = 0.16$), S2 ($p = 1$) and S3 ($p = 1$). Individual data analysis revealed that only 3 participants did not show a decrease in SJ_H at Exh and more than half of the individuals (n = 11) did not recover their performance at $Post_{2min}$ according to the MDC_{95} (Fig. 2B).

3.3 Sit-to-stand performance

A significant time effect was found for STS_T ($p < 0.001$; $\eta_p^2 = 0.73$). STS_T rose progressively along the fatiguing

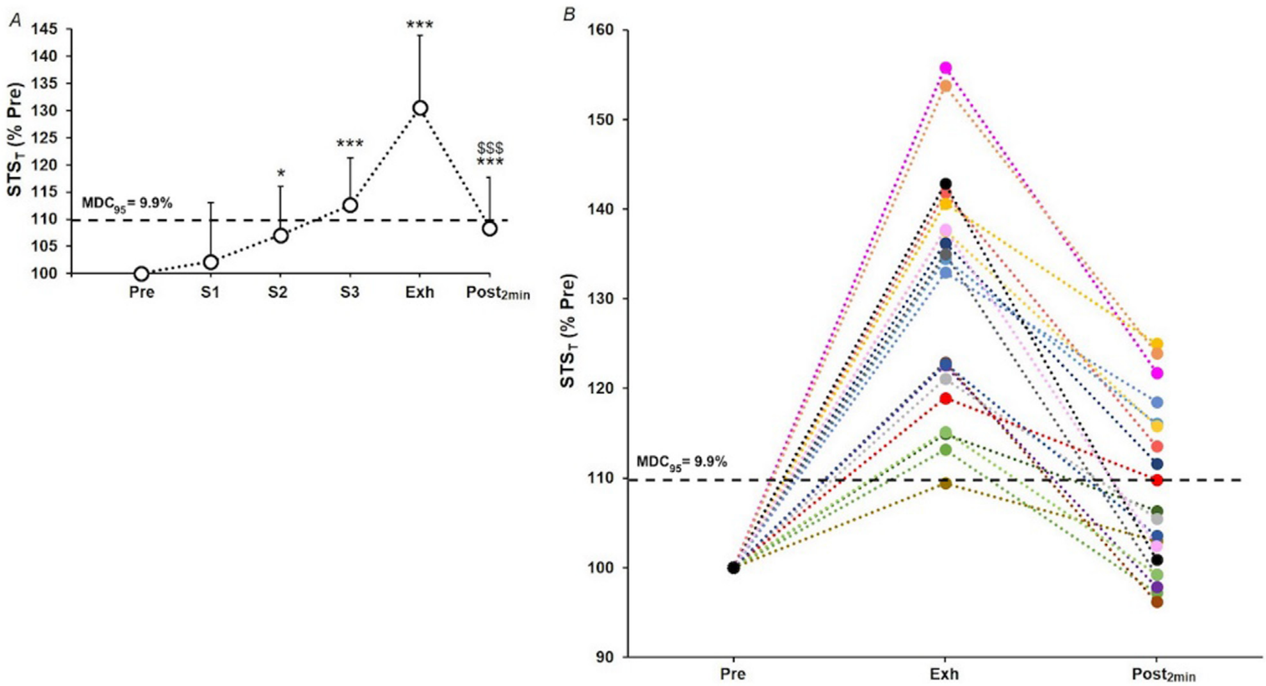

Fig. 3. Sit-to-stand time (STS_T) recorded before (Pre), during (S1, S2, S3), at exhaustion (Exh) and 2 min after the end of the fatiguing test ($Post_{2min}$). Panel A displays the pooled data of the 20 participants. Panel B displays the individual data for values recorded before the protocol (Pre), at exhaustion (Exh) and after a 2-min recovery period ($Post_{2min}$). Error bars on the panel A denote standard deviation. Horizontal dashed line represents the minimal detectable change (MDC_{95}). Data are expressed as percentage of Pre. Significantly different from Pre: * $p < 0.05$, *** $p < 0.001$. Significantly different from Exh: $$$ $p < 0.001$

protocol, with mean increases of 2.2 ± 10.8, 7.0 ± 9.0, 12.6 ± 8.7 and $30.5 \pm 13.3\%$ at S1, S2, S3 and Exh when compared to Pre, respectively (Fig. 3A). This increase in STS_T was non-significant at S1 but became significant from S2 and for all other times measurements ($p < 0.01$) when compared to baseline. Mean of changes in STS_T was greater than the MDC_{95} (0.053 s; 9.96%) only for S3 and Exh. Two minutes of recovery were not sufficient for STS_T to fully recover (significant Pre vs. $Post_{2min}$ differences; $p < 0.001$), however mean decrease in performance at $Post_{2min}$ ($8.4 \pm 9.4\%$) was not greater than MDC_{95}. Individual data analysis revealed that only one participant did not show an increase in STS_T at Exh and less than half of the individuals (n = 8) did not recover their performance at $Post_{2min}$ according to the MDC_{95} (Fig. 3A). Raw data for SJ_H and STS_T recorded at each time measurements are displayed in Table 2.

4 Materials and methods (experiment B)

4.1 Participants

An independent sample of fifteen participants (four females; mean ± SD age: 22 ± 3 years, height: 173 ± 10 cm, weight: 66 ± 11 kg) participated in experiment B. They visited the laboratory on three occasions, i.e., one familiarization and two experimental sessions separated by 3–7 days. The familiarization and experimental sessions were exactly the same to the one of the Experiment A.

4.2 Procedures

Considering the above-presented results, no change in the fatigability protocol was necessary. The fatigability protocol used for experiment B was rigorously the same than in the experiment A (see Fig. 1). For squat jump performance assessment, we ran the *My Jump 2* application on both an iPad mini ® (120 fps) and an iPad Pro ® (240 fps). This was done to compare the test-reliability of each recording mode to ensure that the fatigability outcomes could be reliably assessed with any commercial devices. Since no difference was found in SJ_H between recording modes at any time point (i.e., Pre, S1, S2, S3, Exh, $Post_{2min}$), and excellent reliability ($ICC_{2\text{-}k} \geq 0.99$) was reported, only data obtained from the 240-fps mode are presented. For the evaluation of STS_T, we used the same device as in the Experiment A (i.e., iPhone X, 240 fps; Apple Inc., California, USA) to run the *Sit to Stand* application.

4.3 Statistical analysis

Sample size estimation for Experiment B was based on an expected excellent reliability ($\rho H_1 \geq 0.9$) based on the results of the Experiment A, an alpha error of 0.5, a statistical power of 80% for a two-tailed test, and a minimum acceptable reliability of moderate ($\rho H_0 \geq 0.5$). The estimated sample size was 12 participants.

The same group isotime analysis was performed. The lowest number of completed stages was three (S1, S2 and S3). Differences in fatigability indicators (i.e., SJ_H, STS_T,

Table 2. My Jump- and Sit to Stand-related outcomes recorded during the first three stages of the fatigability test (S1, S2 and S3), at exhaustion (Exh) and 2 min after the end of the fatigability test ($Post_{2min}$)

	Pre	S1	S2	S3	Exh	$Post_{2min}$
SJ_H (m)	0.278 ± 0.07	0.26 ± 0.07	0.248 ± 0.07	0.244 ± 0.07	0.217 ± 0.06	0.247 ± 0.07
STS_T (s)	0.536 ± 0.06	0.547 ± 0.86	0.575 ± 0.09	0.604 ± 0.09	0.699 ± 0.01	0.58 ± 0.08

Values are mean ± SD. SJ_H: squat jump height; STS_T: sit-to-stand time.

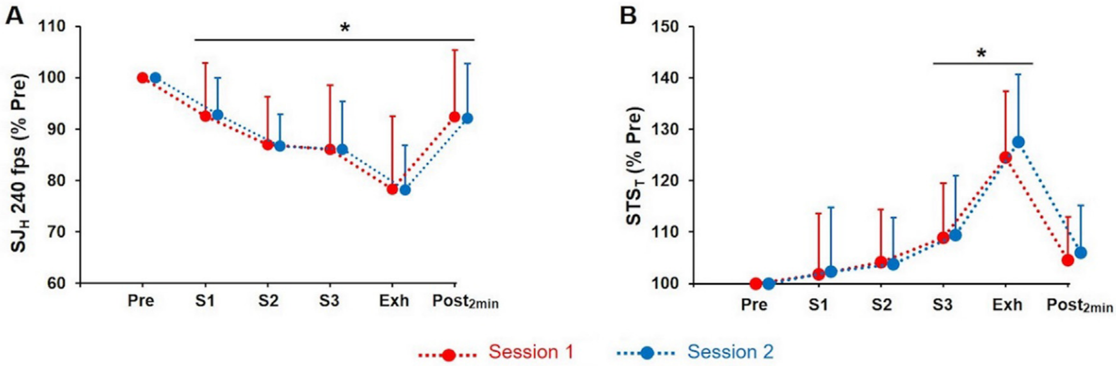

Fig. 4. Squat jump height (SJ_H) recorded at 240 per second (panel A), and sit-to-stand time (STS_T; panel B) recorded before (Pre), during (S1, S2, S3), at exhaustion (Exh) and 2 min after ($Post_{2min}$) the end of the fatiguing protocols performed in two separate experimental sessions. Error bars denote standard deviation. Data are expressed as percentage of Pre. Significantly different from Pre: * $p < 0.05$

HR, RPE_M, RPE_R) between sessions were tested using two-way repeated measures ANOVAs [time (i.e., Pre, S1, S2, S3, Exh, $Post_{2min}$) × session (session 1, session 2)]. If significant main effects were observed, these were followed up by post hoc Bonferroni-corrected pairwise comparisons. Finally, for test-retest reliability analyses of fatigability indicators, the intra-class correlation coefficient two-way mixed effects consistency ($ICC_{3\text{-}1}$) was used as a measure of relative reliability and interpreted as poor (<0.5), moderate (0.50–0.74), good (0.75–0.89), and excellent (≥0.9). Absolute reliability was assessed using typical error expressed as a coefficient of variation (CV_{TE}) which was calculated using the online spreadsheet of Hopkins (Hopkins, 2000). The alpha level for all statistical tests was set a $p < 0.05$.

5 Results (experiment B)

5.1 Endurance time, heart rate and rating of perceived exertion

The mean number of completed stages was 4.9 ± 1.4 (range 3–8 stages) and 5.0 ± 1.4 (range 3–8 stages) in sessions 1 and 2, respectively, with a mean exercise duration of 320 ± 169 (range 134–709 s) and 324 ± 175 s (range 111–704 s). For exercise duration, CV_{TE} between both sessions was 19.4% and $ICC_{3\text{-}1}$ was 0.87 (95% CI: 0.65–0.96). A significant effect of time ($p < 0.001$; $\eta_p^2 = 0.85$) but no time × session interaction ($p = 0.08$; $\eta_p^2 = 0.15$) was found for HR. The mean HR_{PEAK} obtained during the fatigability protocols was 86.3 ± 5.6 and 88.4 ± 5.6% of theoretical HR_{MAX} for sessions 1 and 2, respectively. For this variable, CV_{TE} was 2.0% and $ICC_{3\text{-}1}$ was 0.90 (95% CI: 0.73–0.97). While a main effect of time was found for both RPE_R and RPE_M (both $p < 0.001$, with values recorded at S2, S3 and Exh being significantly higher than S1), no time × session interaction was found for these two variables ($p = 0.19$ and 0.90, respectively). RPE_M and RPE_R recorded at Exh during the first and second sessions were 19.8 ± 0.6 and 13.8 ± 0.3, and 19.8 ± 0.4 and 13.7 ± 2.9, respectively ($p > 0.05$). $ICCs_{3\text{-}1}$ and CV_{TE} were 0.79 (95% CI: 0.49–0.93) and 9.4% for RPE_M and 0.5 (95% CI: 0.03–0.8) and 17.4% for RPE_R, both being obtained at isotime (i.e., stage 3).

5.2 Squat jump and sit-to-stand performance

A significant main effect of time was found for SJ_H ($p < 0.001$; $\eta_p^2 = 0.58$), with each time measurements being significantly different when compared to Pre (Fig. 4A), and a significant time effect was found for STS_T ($p < 0.001$; $\eta_p^2 = 0.72$), with values recorded at S3 and Exh being significantly different from Pre (Fig. 4B). No time × session for SJ_H and STS_T interactions were found ($p > 0.05$). The reliability of SJ_H and STS_T at isotime (i.e., S3), exhaustion and during the recovery period is displayed in Table 3.

6 Discussion

Two distinct experiments were conducted to develop an easy-to-implement fatigability test and assess its test-retest reliability. Progressive development of muscle

Table 3. Between sessions reliability for squat jump height (SJ_H) and sit-to-stand time (STS_T) at isotime (i.e. stage 3), immediately after task failure (Exh) and after two min of recovery ($Post_{2min}$).

	Session 1	Session 2	Change in mean (95% CI)	CV_{TE} (95% CI)	$ICC_{3\text{-}1}$ (95% CI)
Isotime					
SJ_H (m)	0.238 ± 0.08	0.237 ± 0.09	0.0002 (–0.012 to 0.012)	6.6 (4.8 to 10.4)	0.97 (0.91 to 0.99)
STS_T (s)	0.569 ± 0.08	0.569 ± 0.08	–0.0006 (–0.007 to 0.006)	1.4 (1 to 2.2)	0.99 (0.97 to 0.99)
Exh					
SJ_H (m)	0.214 ± 0.08	0.217 ± 0.09	–0.0012 (–0.013 to 0.011)	7.4 (5.4 to 11.7)	0.96 (0.89 to 0.99)
STS_T (s)	0.65 ± 0.09	0.66 ± 0.08	–0.011 (–0.03 to 0.005)	3.2 (2.4 to 5.1)	0.93 (0.81 to 0.97)
$\mathbf{Post_{2min}}$					
SJ_H (m)	0.253 ± 0.09	0.255 ± 0.09	–0.0019 (–0.015 to 0.012)	6.8 (5 to 10.7)	0.97 (0.91 to 0.99)
STS_T (s)	0.54 ± 0.06	0.55 ± 0.07	–0.005 (–0.01 to 0.005)	2.4 (1.7 to 3.8)	0.96 (0.88 to 0.98)

Values are mean ± SD. 95% CI: 95% confidence interval.

fatigability was detected during the test performed in the first experiment. The second experiment reported that the test-retest reliability was excellent for both SJ_H and STS_T changes during and after the test.

6.1 Test ability to detect muscle fatigability

Our first objective was to evaluate if our test conception was effective to induce progressive muscle fatigability. Several studies reported a significant level of muscle fatigability when the indicator of muscle performance changed by ~10% compared to baseline (Bachasson, Guinot *et al.*, 2013; Doyle-Baker et al., 2017; Souron *et al.*, 2020). When looking at individual data, SJ_H was decreased by > 10% at Exh in 16/20 participants, and STS_T was increased > 10% in 19/20 participants. This fatigability threshold of 10% is, however, arbitrary. Based on MDC_{95} analysis, we further reported that only three and one participants did not show significant changes in SJ_H and STS_T at Exh, respectively. This confirms that our test induced important levels of muscle fatigability in almost all individuals. Importantly, muscle fatigability can be easily and affordably detected through two validated mobile applications highlighting its potential to be used in clinical settings where space and technology are often constrained. Monitoring the decrease in performance due to muscle fatigability during functional tasks such as rising from a chair or jumping could be more informative of its impact on daily living activities or athletic performance than monitoring muscle fatigability through isolated isometric contraction.

6.2 Test-retest reliability

Excellent reliability was found for SJ_H and STS_T (see Table 3). The reliability results of our test are nearly comparable with those of similar studies that developed more complex test using intermittent isometric (e.g., CV_{TE} of 5.3% and ICC of 0.88 in Bachasson *et al.*, 2013) and dynamic (e.g., ICC of 0.84–0.89 in Maffiuletti, Bizzini, Desbrosses, Babault and Munzinger (2007)) muscle contractions, but with the advantage to be portable and not to require any technical expertise. This is promising for the future dissemination of our test.

The total exercise duration exhibited a good relative ($ICC_{3\text{-}1}$ of 0.87) but only moderate absolute reliability (CV_{TE} of 19.4%). These results are better than other timed wall squat tests (Lubans *et al.*, 2011) but slightly lower when compared to submaximal isometric tests performed until exhaustion (Bachasson, Millet *et al.*, 2013; Gruet et al., 2010; Gruet, Vallier, Mely, & Brisswalter, 2010). Several arguments can explain this moderate absolute reliability. Endurance performance is influenced by external motivation (Marcora & Staiano, 2010). This is especially true for long-duration exercise modalities, but also for "open ended"-exhaustion tasks compared to time-trial where the endpoint is known (Jeukendrup, Saris, Brouns, & Kester, 1996). We decided not to provide strong encouragements during this test, as the frequency and intensity of the encouragements are not easy to standardize and replicate. Alternatively, we decided to use very simple sentences pronounced with a neutral tone at regular time intervals. We believe this procedure may reduce variability across testing locations, but it is still possible that the absence of uninterrupted strong encouragements close to exhaustion may have slightly increased intra-individual variability in the specific context of a reliability study, where participants performed the test several times in the absence of a highly motivational purpose. The lower reliability of endurance performance compared to indicators of fatigability further emphasizes the value of the latter. These indicators possess the advantage of being measured using valid, affordable and portable instruments as early as the first stages and throughout the exercise duration, with time-points far to exhaustion being much less influenced by the potential motivation bias.

The test-retest reliability of perceptual measurements obtained at isotime was satisfactory for muscular (RPE_M) but not respiratory (RPE_R) efforts. Respiratory effort was only moderate at isotime (i.e., 9.9 ± 2.3). Participants likely experienced more difficulties in reporting RPE_R as they were less accustomed to focus their attention on respiratory sensations which have a minor influence on the overall sensory integration process during activities of daily living. This may explain the lower test-retest reliability of RPE_R compared to RPE_M. We may anticipate that respiratory effort would be more important in specific populations (e.g., people with chronic respiratory disorders) that are more accustomed to these sensations in their everyday life, potentially leading to better reliability in RPE_R. The reliability of RPE_R can be important if our test is intended to be used in specific populations in the future, but we could argue that it is not useful to monitor this parameter when using this test in the general population, and that only RPE_M should be collected and considered for further analyses.

6.3 Test feasibility

The feasibility of our test was excellent, with no adverse events reported in any participants. We did not evaluate the presence of muscle damage in the days that followed the test. No differences were found in baseline SJ_H and STS_T between the first and second sessions of the Experiment B, suggesting that the potential damages of the contractile apparatus have recovered before the second experimental session, i.e. at least 3 days. This has important implications as exercise-induced muscle damage may be disruptive for the engagement in subsequent physical activity, especially in frail populations, potentially decreasing the uptake of a test. The full protocol has an acceptable duration of < 30 min, and requires only little and unexpensive equipment, contrasting with classical time-consuming and less ecological muscle fatigability test (Bachasson, Millet *et al.*, 2013; Maffiuletti *et al.*, 2007; Pageaux *et al.*, 2016 Pageaux, Lepers, & Marcora, 2016). The intermittent nature of our test together with its low exercise duration and submaximal cardiac stress further reinforce its feasibility for some specific populations who are sometimes unable to sustain prolonged periods of continuous exercise (Beauchamp *et al.*, 2010). Sustained isometric contraction may lead to muscle ischemia due to increased intramuscular tissue pressure (Sadamoto, Bonde-Petersen, & Suzuki, 1983), which could be viewed as a factor limiting the feasibility of isometric wall-squat test in some clinical populations. However, it should be noted that such isometric modalities have been employed in various populations (e.g. people with hypertension) for both testing and training purposes, with high level of feasibility and safety (Baffour-Awuah *et al.*, 2023; Wiles, Goldring, O'Driscoll, Taylor, & Coleman, 2018). Altogether, these arguments support the potential of this test for being used in a large variety of populations (e.g. athletes, people with specific medical conditions), which remain to be confirmed by future studies investigating its reliability and feasibility in specific contexts.

6.4 Practical application

This test could be used in applied settings as a standard exercise test or a screening instrument, similar to other simple tests such as walking or sit-to-stand. These field tests have the drawback of being self-paced, with performance potentially influenced by pacing strategies. These tests may also suffer from a ceiling effect, limiting their use in the general population and even in people with chronic disease (Radtke, Puhan, Hebestreit, & Kriemler, 2016). Our test overcomes these limitations, and also has the advantage of cumulating two different indicators of muscle performance. This offers flexibility based on the population being studied. Additionally, this test can be used to assess the efficacy of a given acute and/or chronic intervention on muscle fatigability. The ease of our test favours multicentre studies, which is a real asset compared to most studies having muscle fatigability as a primary outcome which are often characterized by small sample size. Our test is not designed as most other fatigability tests used in laboratory settings that are generally based on a percentage of a maximal value (e.g., muscle strength). This experimental approach does not reflect ecologically valid conditions where the mobilized muscle mass and the needed muscle force are imposed by the task. Our test overcomes this limitation with the intensity of exercise reflecting an absolute quantity of muscle work, that is influenced by the participant' body mass.

7 Conclusion

We developed a new muscle fatigability test which demonstrates an excellent feasibility and test-retest reliability. The low resources requirements and easy-to-implement nature of this test make it a promising tool for evaluating muscle fatigability in both applied and laboratory settings. Future studies should i) confirm its feasibility and reliability in clinical populations and ii) focus on the establishment of reference values in a wide range of age groups, ensuring optimal interpretation in specific populations.

Funding

This work was supported by a grant from "Toulon, Provence, Méditerranée" (N° 201445)

Data availability statement

Data generated or analysed during this study are provided in full within the published article.

Conflict of interests

The authors declare there are no competing interests.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. The CERSTAPS research ethic committee (n° IRB00012476-2021-08-02-87) approved all procedures. All participants provided written informed consent.

Remerciements

Ce travail a bénéficié d'une subvention de la part de "Toulon, Provence, Méditerranée" (N° 201445).

Le matériel de l'article a fait l'objet d'une présentation lors du congrès de l'Association des Chercheurs en Activités Physiques et Sportives (ACAPS) 2021, à Montpellier.

References

Ayotte, N. W., Stetts, D. M., Keenan, G., & Greenway, E. H. (2007). Electromyographical analysis of selected lower extremity muscles during 5 unilateral weight-bearing exercises. *Journal of Orthopaedic & Sports Physical Therapy*, *37* (2),48–55. doi: 10.2519/jospt.2007.2354

Bachasson, D., Guinot, M., Wuyam, B., Favre-Juvin, A., Millet, G. Y., Levy, P., & Verges, S. (2013). Neuromuscular fatigue and exercise capacity in fibromyalgia syndrome. *Arthritis Care & Research*, *65* (3), 432–440. doi: 10.1002/acr.21845

Bachasson, D., Millet, G. Y., Decorte, N., Wuyam, B., Levy, P., & Verges, S. (2013). Quadriceps function assessment using an incremental test and magnetic neurostimulation: A reliability study. *Journal of Electromyography and Kinesiology*, *23* (3), 649–658. https://doi.org/10.1016/j.jelekin.2012.11.011

Baffour-Awuah, B., Pearson, M. J., Dieberg, G., Wiles, J. D., & Smart, N. A. (2023). An evidence-based guide to the efficacy and safety of isometric resistance training in hypertension and clinical implications. *Clinical Hypertension*, *29* (1), 9 doi: 10.1186/s40885-022-00232-3

Balsalobre-Fernandez, C., Glaister, M., & Lockey, R. A. (2015). The validity and reliability of an iPhone app for measuring vertical jump performance. *Journal of Sport Sciences*, *33* (15), 1574–1579. https://doi.org/10.1080/02640414.2014.996184

Beauchamp, M. K., Nonoyama, M., Goldstein, R. S., Hill, K., Dolmage, T. E., Mathur, S., & Brooks, D. (2010). Interval versus continuous training in individuals with chronic obstructive pulmonary disease-a systematic review. *Thorax*, *65* (2), 157–164. doi: 10.1136/thx.2009.123000

Bolgar, M. R., Baker, C. E., Goss, F. L., Nagle, E., & Robertson, R. J. (2010). Effect of exercise intensity on differentiated and undifferentiated ratings of perceived exertion during cycle and treadmill exercise in recreationally active and trained women. *Journal of Sports Science and Medicine*, *9* (4), 557.

Borg, G. (1998). Borg's perceived exertion and pain scales. *Champaing, IL: Human kinetics*, *104*.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, *100* (2), 126.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Erlbaum. *Hillsdale, NJ*

De Pauw, K., Roelands, B., Cheung, S. S., de Geus, B., Rietjens, G., & Meeusen, R. (2013). Guidelines to classify subject groups in sport-science research. *International Journal of Sports Physiology and Performance*, *8* (2), 111–122. https://doi.org/10.1123/ijspp.8.2.111

Doyle-Baker, D., Temesi, J., Medysky, M. E., Holash, R. J., & Millet, G. Y. (2017). An Innovative Ergometer to Measure Neuromuscular Fatigue Immediately after Cycling. *Medicine & Science in Sports & Exercise*, *50* (2), 375–387 https://doi.org/10.1249/MSS 0000000000001427

Enoka, R. M., & Duchateau, J. (2016). Translating fatigue to human performance. *Medicine & Science in Sports & Exercise*, *48* (11), 2228–2238. https://doi.org/10.1249/MSS 0000000000000929

Evans, R. A., Kaplovitch, E., Beauchamp, M. K., Dolmage, T. E., Goldstein, R. S., Gillies, C. L., Mathur, S. (2015). Is quadriceps endurance reduced in COPD?: a systematic review. *Chest*, *147* (3), 673–684. doi: 10.1378/chest.14-1079

Gellish, R. L., Goslin, B. R., Olson, R. E., McDonald, A., Russi, G. D., & Moudgil, V. K. (2007). Longitudinal modeling of the relationship between age and maximal heart rate. *Medicine & Science in Sports & Exercise*, *39* (5), 822–829. https://doi.org/10.1097/mss.0b013e31803349c6

Goldring, N., Wiles, J. D., & Coleman, D. (2014). The effects of isometric wall squat exercise on heart rate and blood pressure in a normotensive population. *Journal of Sports Sciences*, *32* (2), 129–136. doi: 10.1080/02640414.2013.809471

Green, J., Crews, T., Bosak, A., & Peveler, W. (2003). Overall and differentiated ratings of perceived exertion at the respiratory compensation threshold: effects of gender and mode. *European Journal of Applied Physiology*, *89*, 445–450. doi: 10.1007/s00421-003-0869-4

Gruet, M., Decorte, N., Mely, L., Vallier, J.-M., Camara, B., Quetant, S., Wuyam, B., Verges, S. (2016). Skeletal muscle contractility and fatigability in adults with cystic fibrosis. *Journal of Cystic Fibrosis*, *15* (1), e1–e8. doi: 10.1016/j.jcf.2015.05.004

Gruet, M., Mely, L., & Vallier, J. M. (2018). Overall and differentiated sensory responses to cardiopulmonary exercise test in patients with cystic fibrosis: kinetics and ability to predict peak oxygen uptake. *European Journal of Applied Physiology*, *118* (9), 2007–2019. https://doi.org/10.1007/s00421-018-3923-y

Gruet, M., Vallier, J., Mely, L., & Brisswalter, J. (2010). Long term reliability of EMG measurements in adults with cystic fibrosis. *Journal of Electromyography and Kinesiology*, *20* (2), 305–312. doi: 10.1016/j.jelekin.2009.05.001

Hopkins, W. (2000). Reliability from consecutive pairs of trials. *A new view of statistics*.

Jeukendrup, A., Saris, W. H., Brouns, F., & Kester, A. D. (1996). A new validated endurance performance test. *Medicine and Science in Sports and Exercise*, *28* (2), 266–270.

Lea, J. W., O'Driscoll, J. M., Coleman, D. A., & Wiles, J. D. (2021). Validity and reliability of RPE as a measure of intensity during isometric wall squat exercise. *Journal of Clinical and Translational Research*, *7* (2), 248.

Lubans, D. R., Morgan, P., Callister, R., Plotnikoff, R. C., Eather, N., Riley, N., & Smith, C. J. (2011). Test-retest reliability of a battery of field-based health-related fitness measures for adolescents. *Journal of Sports Sciences*, *29* (7), 685–693. doi: 10.1080/02640414.2010.551215

Maffiuletti, N. A., Bizzini, M., Desbrosses, K., Babault, N., & Munzinger, U. (2007). Reliability of knee extension and flexion measurements using the Con-Trex isokinetic dynamometer. *Clinical Physiology and Functional Imaging*, *27* (6), 346–353. doi: 10.1111/j.1475-097X 2007. 00758.x

Marcora, S. M., & Staiano, W. (2010). The limit to exercise tolerance in humans: mind over muscle? *European Journal of Applied Physiology*, *109* (4), 763–770. https://doi.org/10.1007/s00421-010- 1418-6

Nicolò, A., Sacchetti, M., Girardi, M., McCormick, A., Angius, L., Bazzucchi, I., & Marcora, S. M. (2019). A comparison of different methods to analyse data collected during time-to-exhaustion tests. *Sport Sciences for Health*, *15* (3), 667–679.

Pageaux, B., Lepers, R., & Marcora, S. M. (2016). Reliability of a novel high intensity one leg dynamic exercise protocol to measure muscle endurance. *PLoS One*, *11* (10), e0163979. doi: 10.1371/journal.pone.0163979

Pageaux, B., Marcora, S. M., & Lepers, R. (2013). Prolonged mental exertion does not alter neuromuscular function of the knee extensors. *Medicine & Science in Sports & Exercise*, *45* (12), 2254–2264. https://doi.org/10.1249/MSS.0b013e31829b504a

Radtke, T., Puhan, M. A., Hebestreit, H., & Kriemler, S. (2016). The 1-min sit-to-stand test—A simple functional capacity test in cystic fibrosis? *Journal of Cystic Fibrosis*, *15* (2), 223–226. doi: 10.1016/j.jcf.2015.08.006

Robertson, R. J., & Noble, B. J. (1997). Perception of physical exertion: methods, mediators, and applications. *Exercise and Sport Sciences Reviews*, *25*, 407–452.

Ruiz-Cardenas, J. D., Rodriguez-Juan, J. J., Smart, R. R., Jakobi, J. M., & Jones, G. R. (2018). Validity and reliability of an iPhone App to assess time, velocity and leg power during a sit-to-stand functional performance test. *Gait Posture*, *59*, 261–266. https://doi.org/10.1016/j.gaitpost.2017.10.029

Sadamoto, T., Bonde-Petersen, F., & Suzuki, Y. (1983). Skeletal muscle tension, flow, pressure, and EMG during sustained isometric contractions in humans. *European Journal of Applied Physiology and Occupational Physiology*, *51* (3), 395–408. doi: 10.1007/BF00429076

Souron, R., Voirin, A. C., Kennouche, D., Espeit, L., Millet, G. Y., Rupp, T., & Lapole, T. (2020). Task failure during sustained low-intensity contraction is not associated with a critical amount of central fatigue. *Scandinavian Journal of Medicine & Science in Sports*, *30* (12), 2329–2341. https://doi.org/10.1111/sms.13815

Strassmann, A., Steurer-Stey, C., Lana, K. D., Zoller, M., Turk, A. J., Suter P., & Puhan, M. A. (2013). Population-based reference values for the 1-min sit-to-stand test. *International Journal of Public Health*, *58* (6), 949–953. doi: 10.1007/s00038-013-0504-z

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, *19* (1), 231–240. doi: 10.1519/15184.1

Wiles, J. D., Goldring, N., O'Driscoll, J. M., Taylor, K. A., & Coleman, D. A. (2018). An alternative approach to isometric exercise training prescription for cardiovascular health. *Translational Journal of the American College of Sports Medicine*, *3* (2), 10–18

Wiles, J. D., Taylor, K., Coleman, D., Sharma, R., & O'Driscoll, J. M. (2018). The safety of isometric exercise: Rethinking the exercise prescription paradigm for those with stage 1 hypertension. *Medicine*, *97* (10). doi: 10.1097/MD 0000000000010105