



Diagnostic Methods

Reliability of knee extensor neuromuscular structure and function and functional tests' performance



Fábio J. Lanferdini ^{a,*}, Francesca C. Sonda ^b, Isabel A. Paz ^b, Luma Z. Oliveira ^b,
Edgar S. Wagner Neto ^b, Talita Molinari ^b, Daniele C.S. Gomes ^b, Graciele Sbruzzi ^b,
Marco A. Vaz ^b

^a Laboratory of Biomechanics, Center of Sports, Federal University of Santa Catarina - Florianópolis, Santa Catarina, Brazil

^b Exercise Research Laboratory, School of Physical Education, Physiotherapy and Dance, Federal University of Rio Grande do Sul - Porto Alegre, Rio Grande do Sul, Brazil

ARTICLE INFO

Article history:

Received 7 February 2021

Received in revised form

13 April 2021

Accepted 7 May 2021

Keywords:

Intra-rater

Inter-rater

Inter-analyzer

Healthy subjects

Reliability

ABSTRACT

Introduction: The aim of this study was to evaluate the intra and inter-rater and inter-analyzer reliability of neuromuscular variables and functional tests.

Methods: Cross-sectional crossover design. Two independent raters and analyzers evaluated twenty-two healthy subjects. Knee-extensor strength was assessed from three maximal voluntary isometric contractions. Muscle activation was obtained from the vastus lateralis (VL), rectus femoris (RF), and vastus medialis (VM) muscles. VL and RF muscles' architecture [fascicle length (FL), pennation angle (PA), muscle thickness (MT)] was obtained at rest by ultrasound. The time from five sit-to-stand (STS) trials, and the distance from the 6-min walk test (6MWT) were obtained. Intraclass correlation coefficient was determined and classified as strong ($r = 0.75$ – 1.00), moderate ($r = 0.40$ – 0.74), and weak ($r < 0.40$).

Results: Strong intra-rater reliability values were observed for strength ($r = 0.97$), muscle activation [VL ($r = 0.91$); RF ($r = 0.92$); VM ($r = 0.80$)], VL [FL ($r = 0.90$); PA ($r = 0.94$); MT ($r = 0.99$)] and RF [MT ($r = 0.85$)] muscle architecture, STS ($r = 0.95$), and 6MWT ($r = 0.98$). Inter-rater reliability also presented strong values for strength ($r = 0.97$), muscle activation [VL ($r = 0.94$); RF ($r = 0.79$); VM ($r = 0.78$)], muscle architecture VL [PA ($r = 0.81$) and MT ($r = 0.88$)] and RF [MT ($r = 0.80$)], STS ($r = 0.93$), and 6MWT ($r = 0.98$). A moderate correlation VL muscle architecture [FL ($r = 0.69$)]. Inter-analyzer muscle architecture reliability presented strong VL [FL ($r = 0.77$); PA ($r = 0.76$); MT ($r = 0.91$)] and RF [MT ($r = 0.99$)].

Conclusion: The high intra and inter-rater and inter-analyzer reliability values for most variables is evidence that they can be used for clinical evaluation. Muscle architecture might need a longer training period by different raters and analyzers to increase reliability.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

In clinical practice, the evaluation of intra-rater and inter-rater reliability is extremely important to ascertain possible influence of different days of evaluation and of different raters on the outcome variables' results. In addition, it also allows determining strategies to minimize measurement errors in future evaluation

protocols.

Several studies have determined the reliability of maximum force (Kollock et al., 2010; Lienhard et al., 2013; McCarthy et al., 2008; Ruschel et al., 2015; Suzuki, 2015; Toonstra and Mattacola, 2013), muscle activation (Balshaw et al., 2017; Callaghan et al., 2009; Hashemi Oskouei et al., 2013; Place et al., 2007; Sorbie et al., 2018) and muscle architecture (Ishida et al., 2018; Kwah et al., 2013; Lima et al., 2012; Lima and Oliveira, 2013; Marzilger et al., 2018; Raj et al., 2012; Santos & Armada-da-Silva, 2017; Silva et al., 2018).

For knee extensor force measurements, most studies have found high values for both intra-rater (Kollock et al., 2010; Suzuki, 2015; Toonstra and Mattacola, 2013) and inter-rater (Kollock et al., 2010)

* Corresponding author. Laboratório de Biomecânica, Centro de Desportos, Universidade Federal de Santa Catarina Rua Deputado Antônio Edu Vieira, Campus Universitário - Trindade, Florianópolis, SC, CEP 88040-900, Brazil.

E-mail address: fabiolanferdini@gmail.com (F.J. Lanferdini).

Abbreviation list

ATS	American Thoracic Society
AG	Acceleration of Gravity
AUC	Area of force under the curves
CV	Coefficient of variation
FL	Fascicle length
icc	Intraclass correlation coefficient
TL	Tibia's Length
M	Mean
mdc	Minimum detectable change
mt	Muscle thickness

mvic(s)	Maximal voluntary isometric contractions(s)
pa	Pennation angle
rf	Rectus femoris
RMS	Root mean square
SD	Standard deviation
sem	Standard error of the measurement
sts	Sit-to-stand
vi	Vastus intermedius
vl	Vastus lateralis
vm	Vastus medialis
6mwt	Six-minute walk test

reliability. This high reliability of knee extensor force measurements was evaluated with load cells (intra-rater $r = 0.89–0.96$) (Kollock et al., 2010; Suzuki, 2015; Toonstra and Mattacola, 2013). However, different protocols and evaluation devices were used for these strength measures, and the reliability of the knee extensor strength measures of healthy subjects, evaluated with knee and hip at 90° of flexion, sitting on a standard chair, is still uncertain.

Intra-rater reliability of knee extensor muscles' activation measures has been investigated during maximal (Callaghan et al., 2009; Hashemi Oskoueie et al., 2013; Place et al., 2007) and sub-maximal (Sorbie et al., 2018) voluntary contractions. During maximal voluntary isometric contractions (MVICs) of knee extensors, Place et al. (2007) found high reliability values for the vastus lateralis (VL, $r = 0.78$) and rectus femoris (RF, $r = 0.90$) muscles, whereas moderate reliability values were found for the vastus medialis (VM, $r = 0.45$) muscle.

Some studies have investigated the quadriceps muscle architecture's reliability (Ishida et al., 2018; Lima et al., 2012; Lima and Oliveira, 2013; Marzilger et al., 2018; Santos & Armada-da-Silva, 2017; Silva et al., 2018). The assessment of muscle thickness (MT) of the RF ($r = 0.91–0.99$), VL ($r = 0.80–0.99$), VM ($r = 0.90$) and vastus intermedius (VI, $r = 0.74$) muscles presented high values of intra-rater reliability (Ishida et al., 2018; Lima and Oliveira, 2013; Santos & Armada-da-Silva, 2017). Only one study evaluated the inter-rater reliability of muscle architecture measures, and found high ($r = 0.96$) RF reliability for the MT (Ishida et al., 2018). Reliability evaluation of the muscle's pennation angle (PA) and fascicle length (FL) measurements was investigated by Lima and Oliveira (2013), who found high intra-rater reliability values for PA ($r = 0.83–0.99$) and FL ($r = 0.80–0.99$) of VL. However, no studies were found that evaluated inter-rater reliability for PA or FL in healthy subjects. Marzilger et al. (2018) developed a semi-automated algorithm routine to analyze architectural parameters of the VL muscle. The authors found high intra-day reliability values for MT ($r = 0.96–1.00$) and FL ($r = 0.87–0.98$), and moderate-to-high for PA ($r = 0.58–0.94$). Therefore, despite of the several studies that evaluated the quadriceps muscle architecture, it is still not very clear in the literature the intra and inter-rater reliability of PA and FL measures.

In addition to neural, structural and mechanical variables, investigation of functional tests' reliability, such as the 6-min walk test (6MWT) and sit-to-stand (STS) is clinically important, as these functional tests are often used to evaluate the patients' health conditions (Tveter et al., 2014). In the case of the 6MWT, the American Thoracic Society (ATS) published a guideline (with materials, procedures and verbal commands) on how to conduct the 6MWT (Crapo et al., 2002). Therefore, several studies have sought to investigate the reliability of 6MWT (Bellet et al., 2012; Guerra-Balic et al., 2015; Tveter et al., 2014) and STS (Goldberg et al.,

2012; Schurr et al., 2012; Wang et al., 2012) measurements in patients with various disorders. These studies have found moderate-to-high values for the intra and inter-rater reliability of 6MWT and STS functional tests (Bohannon, 2011; Paul and Canning, 2014; Tveter et al., 2014). However, the results vary greatly among studies due to the different characteristics of each evaluated population.

Therefore, the first aim of the present study was to evaluate the intra and inter-rater reliability of the knee extensor muscles' force and muscular activation. The second aim of the study was to evaluate the intra and inter-rater reliability of the VL and RF muscular architecture parameters. Our third goal was to evaluate the intra and inter-rater reliability of 6MWT and STS tests' performance, as well as to evaluate the reliability of the area under the force curve during STS when performed over a force plate. Finally, we chose raters with two different time experience in using the different measures to check for experience as an intervening factor for reliability.

2. Methods

2.1. Participants

This study is characterized as a cross-sectional with reliability measurements. The sample consisted of twenty-two healthy, physically active subjects. Participants were selected from undergraduate and graduate students from the university. The sample was calculated from the G*Power 3.1.3 software (Fraunhofer Universität Kiel, Germany), with an "Effect Size" of 0.30, $\alpha = 0.05$, and power of 0.80. Subjects were carefully informed about the design of the study and the possible risks and discomforts related to the measurements and then all the subjects signed an informed consent form to participate in the study. This study was conducted according the declaration of Helsinki, and all procedures were approved by the local Institutional Research Ethics Committee (project number 3.458.124). Inclusion criteria included age between 18 and 40 years old, being physically active, do not present any medical restriction on the performance of maximum tests. Exclusion criteria included any history of neuromuscular, metabolic, hormonal, and cardiovascular diseases. Subjects were not taking any medication, which could influence on hormonal and neuromuscular metabolism.

2.2. Data acquisition and analysis

The evaluation protocols were carried out at the exercise research laboratory of the institution where the project was approved. All knee extensor strength and muscle activation tests were performed by two independent raters (12 years of experience = most experienced; 2 months = novice) on the first

evaluation day, and only the most experienced rater repeated the evaluation after one-week interval (Santos et al., 2013), adopting the same procedures previously performed. Similarly, VL muscle architecture (most experienced = 3 years; novice = 3 months), and the clinical 6MWT and STS tests (most experienced = 2 years; novice = 3 months) were evaluated by two different raters in a similar way. Standard operating procedures were developed to systematize the procedures for data collection among the raters (Gough and Hamrell, 2010). All data collections were performed in one month.

Prior to the strength tests, the participants' tibial length (TL), from the dominant limb, was measured for torque calculation, which was performed using the equation (1):

$$\text{Torque} = \text{Force} \cdot \text{AG} \cdot (\text{TL}) \quad (1)$$

Equation 1. Force (kg); AG – Acceleration of Gravity (m/s²); TL – Tibia's Length (m)

After that, participants were asked to sit in a standard chair, with their hips and knees flexed at 90° (0° = full knee extension). Next, participants were asked to produce the maximal knee extensor effort against a Velcro strap that was attached in one extremity to the lower end of the leg (distal end of the tibia), and to a load cell (BTS 200 kg, Porto Alegre, Brazil) on the other end. The load cell was attached to a chain that was fixed to the chair's rear legs. After the leg fixation, a verbal command was given to all participants to perform three 5-sec knee extensor maximal voluntary isometric contractions (MVICs). A 2-min interval was adopted between contractions to avoid possible fatigue effects. The highest MVIC force and torque values were selected for further analysis.

Muscle activation was evaluated using a 4-channel electromyography (EMG) system (EMG 800c, EMG System, São Paulo, Brazil; sampling frequency of 2000 Hz per channel) connected to a notebook computer, along with muscle strength data during the knee extensor MVIC tests (Merletti et al., 2009). The muscles monitored in the lower limbs were RF, VL and VM. The EMG activity of the RF, VL and VM muscles was obtained by pairs of surface electrodes (Ag/AgCl, diameter = 22 mm, Kendall Meditrace 100, Chicopee, Canada; impedance input > 1015 Ω, common bounce rate de –92 dB), which were placed in the muscle belly, parallel to the muscle fibers (Merletti et al., 2009).

The EMG signals were processed using mathematical routines in Matlab® software (MathWorks, Natick, USA). First, the signals were filtered with a third order Butterworth pass-band filter, with cutoff frequencies of 20 and 500 Hz. After filtration, the root mean square (RMS) values from each muscle (RF, VL and VM) were calculated from the signals of the plateau from the 5-sec MVICs' force signals.

An ultrasound system (VIVID I®, GE Healthcare, Waukesha, USA) with a linear array probe (50 mm, 3–10 MHz; L12-3, VIVID I®, GE Healthcare, Waukesha, USA), was used to determine the FL, PA and MT of VL and MT of RF. Ultrasonography images were acquired at rest, with the subjects lying in dorsal decubitus on a stretcher, with the lower limbs fully extended. Three VL and RF images were obtained by each rater (Baroni et al., 2013; Guilhem et al., 2013).

Ultrasound images were analyzed off-line using ImageJ software (National Institute of Health, Bethesda, USA) to measure FL, PA and MT. The analysis of muscle architecture data was performed by two independent analyzers with similar experience (6 months), analyzed three similar images of each muscle, in order to ascertain possible influences of the analyzers on the ultrasound measurements' reliability.

2.3. Functional capacity

Functional capacity was assessed using the STS test. The test

consisted of two attempts while performing five repetitions of the movement of getting up from and sitting on a chair. The time to perform these five repetitions was obtained with a stopwatch (CASIO HS-30W-N1V, Computer Co., Ltd. Tokyo, Japan). A 2-min interval was adopted between attempts. The test was performed in a chair with back support, so that, at the beginning of the test, the hip and knee joints were positioned at a 90° angle (0° = full extension). The participants started the test with their trunk straight and their hands crossed in front of the chest. During the test, the participants' feet were maintained apart at approximately the shoulders' width, and above a force plate (9260AA3; Kistler Instrument AG, Winterthur, Suisse). Besides the time, the areas of the ground reaction force curves (N·s) were obtained from the 1st, 3rd and 5th repetitions using MATLAB® software (MathWorks, Natick, USA). Force data were filtered with a low-pass Butterworth digital filter, with the cut-off frequency calculated using Winter residue analysis (Winter 2009). After the filtering procedure, the integral of each execution of interest was calculated.

The 6MWT tests were performed according to the ATS recommendations, which published a guideline (with materials, procedures and verbal commands) on how to conduct the 6MWT (Crapo et al., 2002). The subjects were then instructed to walk as fast as possible for 6 min, always with one foot supported on the ground (no bipodal phase). The test was performed on a flat sidewalk, where two cones were placed at a distance of 30 m between them. A stopwatch (CASIO HS-30W-N1V, Computer Co., Ltd. Tokyo, Japan) was used to control the test time, and the verbal commands were given as proposed by ATS (Crapo et al., 2002). The distance was calculated from the number of turns performed during the 6 min, multiplied by 60 m (distance of one lap), and added to the surplus distance performed by the subjects. This distance was measured using a metric fiberglass tape (Vonder, Grupo OVD, Novo Hamburgo, Brazil), with 20 m of length and 1 mm precision.

2.4. Statistical analysis

The data statistical analysis included mean values (M), standard deviation (SD) values, and coefficient of variation (CV) of all data. The data normality was verified by the Shapiro-Wilk test. All data obtained by the same rater at the two different days were used to verify the intra-rater reliability of the outcome measures. In addition, in order to verify the inter-rater reliability of the outcome measures, all the data collected on the first day by the two raters was used. To verify the outcome measures' reliability, the intraclass correlation coefficient (ICC), standard error of the measurement (SEM) and minimum detectable change (MDC) were calculated. The ICC was classified as excellent ($r > 0.90$); good ($r = 0.75–0.90$); moderate ($r = 0.50–0.75$) or poor ($r < 0.50$) according to the literature (Koo and Li, 2016). SEM was estimated using equation: $\text{SEM} = \text{SD} \cdot \sqrt{(1-\text{ICC})}$ described by Weir (2005). The MDC was estimated based on a 95% confidence interval (95%CI), where $\text{MDC} = 1.96 \cdot \text{SEM}$ (Schwenk et al., 2012). The level of significance adopted for all analyzes was set at 5%. All statistical procedures were performed using the statistical package SPSS 20.0 (IBM, Chicago, USA) for Windows.

3. Results

Twenty-two subjects (11♀ e 11♂), with age: 27 ± 5 years; body mass: 69.7 ± 10.6 kg; and stature: 170 ± 10 cm agreed to participate in the study. A high reliability was observed for most of the outcome variables. Results of the knee extensor MVICs' force presented strong ICC values for both intra-rater ($r = 0.970$, $p < 0.001$) and inter-rater ($r = 0.967$, $p < 0.001$) reliabilities. The knee extensor torque results also presented strong ICC values for both the intra-

rater ($r = 0.971$, $p < 0.001$) and inter-rater ($r = 0.968$, $p < 0.001$) reliability. In addition, the force and torque results showed low values of SEM and MDC (Table 1).

The knee-extensor muscle activation reliability was high for both the intra-rater and inter-rater measurements for all evaluated muscles. ICC values were classified as strong in all analyzed correlations ($r > 0.780$, $p < 0.001$). SEM and MDC values were small for the RF and VL muscles, whereas for the VM muscle, values were slightly higher (Table 1).

The VL muscle architecture reliability results presented a strong reliability for the PA measurements for both the intra-rater ($r = 0.940$, $p < 0.001$) and inter-rater ($r = 0.801$, $p < 0.001$) comparisons. MT measurements also showed strong reliability for the intra-rater ($r = 0.987$, $p < 0.001$) and inter-rater ($r = 0.882$, $p < 0.001$) measures. There was also a strong intra-rater reliability ($r = 0.898$; $p < 0.001$) for FL, but for the inter-rater reliability, moderate values were found ($r = 0.688$; $p = 0.005$). Regarding the SEM and MDC results, the intra-rater data presented low values, whereas, for the inter-rater results, higher values of SEM and MDC were found. Furthermore, RF muscle architecture reliability results presented a strong reliability for the MT measurements for both the intra-rater ($r = 0.848$, $p < 0.001$) and inter-rater ($r = 0.803$, $p < 0.001$); Table 2.

For the functional test results, strong intra and inter-rater reliability were also found. ICC values for STS performance were classified as strong for both the intra-rater ($r = 0.950$, $p < 0.001$) and inter-rater ($r = 0.926$, $p < 0.001$) results, and the STS values of SEM and MDC were small for the intra and inter-rater results (Table 3). In addition, the STS results for the 1st, 3rd and 5th repetitions AUC mean values showed strong ICC values for the intra-rater ($r = 0.852$, $p < 0.001$) and inter-rater ($r = 0.801$, $p < 0.001$) results. However, when analyzing the individual curves, it can be observed that the ICCs fall to moderate in the 5th STS repetition for intra ($r = 0.669$; $p = 0.007$) and inter-rater ($r = 0.586$; $p = 0.025$) results.

Corroborating with STS results, the performance of 6MWT also presented strong ICC values for the intra-rater ($r = 0.977$, $p < 0.001$) and inter-rater ($r = 0.977$, $p < 0.001$) comparisons, in addition to presenting low SEM and MDC values (Table 3).

Finally, for the muscle architecture measures inter-analyzer reliability, strong ICC values were found especially for MT of VL ($r = 0.907$; $p < 0.001$) and RF ($r = 0.990$; $p < 0.001$), as well as strong ICC values for FL analyzes ($r = 0.769$; $p = 0.001$) and PA ($r = 0.763$; $p = 0.001$) of VL muscle architecture. However, these last two variables presented higher SEM and MDC values compared to MT (Table 4).

4. Discussion

The results of the present study demonstrated that measures of strength, torque and muscle activation of the knee extensor muscles have a high reliability. These results also demonstrated that even an inexperienced evaluator was able to replicate evaluations similar to the more experienced evaluator. Therefore, these results corroborate with other results found in the literature, which also evaluated the reliability of measures of knee extensor muscles' strength and activation during MVICs (Balshaw et al., 2017; Callaghan et al., 2009) and during submaximal contractions (Hashemi Oskouei et al., 2013; Sorbie et al., 2018).

The reliability results of muscle architecture measurements demonstrated high intra-rater reliability for all measurements. However, greater variability of the data was found in the inter-rater results, which may be associated with the novice experience of the second rater (3 months). Therefore, from these results it is evident that less experienced raters may be a complication for muscle architectural assessments, especially in populations where muscle quality will be worse than in healthy subjects, such as in elderly and clinical patients (Fragala et al., 2015). An alternative to minimize measurement errors is to allow a longer training period for data collection of these variables for the raters. Another alternative is to always keep the same experienced evaluator in all evaluations, thus avoiding greater measurement errors of these variables (Kwah et al., 2013; Silva et al., 2018).

Although the reliability results of inter-analyzer muscle architecture measurements presented strong correction results, FL and PA presented large SEM values. Therefore, as in the inter-rater results, the less experienced analyzer may have influenced the data analysis, and, therefore, a longer training period of the analyzes of this technique is necessary for a greater muscle architecture reliability (Kwah et al., 2013). Another alternative for the adequacy of the muscle architecture analysis process is to automatize the process, through mathematical routines described in the literature (Marzilger et al., 2018). Therefore, we suggest that only experienced evaluators and analysts are used in the evaluation and analysis of the muscle architecture in elderly participants or in patients.

The STS high reliability results, for both intra and inter-rater measurements, agree with previous results found in the literature that investigated the STS measure's reliability with five repetitions performance (Bohannon, 2011; Eden et al., 2018; Goldberg et al., 2012; Mong et al., 2010; Puthoff and Saskowski, 2013). However, the reliability evaluation of the AUC during the STS showed high reliability values only in the first and third repetitions, whereas in the fifth, the values were moderate. These results may be associated

Table 1

Intra and inter-rater reliability results for force, torque, and muscle activation for the rectus femoris (RF), vastus lateralis (VL) and vastus medialis (VM) muscles of the dominant inferior limb. Values of mean \pm standard deviation (SD), intraclass correlation coefficient (ICC), 95% confidence interval (95%CI), p-value, standard error of the measurement (SEM), minimum detectable change (MDC) and coefficient of variation (CV).

	Mean \pm SD	ICC	95%CI	p-value	SEM	MDC	CV (%)
Force (kg) - Intra-Rater	59.2 \pm 20.5	0.970	0.928–0.988	<0.001	3.40	6.66	33.8
Force (kg) - Inter-Rater	57.3 \pm 18.3	0.967	0.920–0.986	<0.001	3.36	6.59	32.2
Torque (N·m) - Intra-Rater	238.1 \pm 81.5	0.971	0.930–0.988	<0.001	13.88	27.21	34.2
Torque (N·m) - Inter-Rater	231.2 \pm 76.6	0.968	0.922–0.987	<0.001	13.70	26.85	33.1
RMS VL (mV) - Intra-Rater	308.3 \pm 118.3	0.905	0.772–0.961	<0.001	36.46	71.46	38.4
RSM VL (mV) - Inter-Rater	306.9 \pm 109.4	0.936	0.845–0.973	<0.001	27.67	54.23	35.6
RMS RF (mV) - Intra-Rater	327.6 \pm 113.9	0.918	0.801–0.966	<0.001	32.62	63.94	34.8
RMS RF (mV) - Inter-Rater	335.5 \pm 105.2	0.786	0.484–0.911	<0.001	27.67	95.41	31.4
RMS VM (mV) - Intra-Rater	287.0 \pm 107.5	0.799	0.515–0.916	<0.001	48.20	94.47	37.5
RMS VM (mV) - Inter-Rater	279.2 \pm 95.2	0.784	0.479–0.910	<0.001	44.24	86.72	34.1

RMS: root mean square.

Table 2

Intra and inter-rater muscle architecture reliability results of the vastus lateralis (VL) and rectus femoris (RF) muscles of the dominant inferior limb. Values of mean \pm standard deviation (SD), intraclass correlation coefficient (ICC), 95% confidence interval (95%CI), p-value, standard error of the measurement (SEM), minimum detectable change (MDC) and coefficient of variation (CV).

	Mean \pm SD	ICC	95%CI	p-value	SEM	MDC	CV (%)
FL - VL (cm) - Intra-Rater	6.57 \pm 0.99	0.898	0.754–0.958	<0.001	0.32	0.62	15.1
FL - VL (cm) - Inter-Rater	6.51 \pm 0.87	0.688	0.250–0.871	0.005	0.49	0.95	13.4
PA - VL (°) - Intra-Rater	20.44 \pm 2.49	0.940	0.856–0.975	<0.001	0.61	1.19	12.2
PA - VL (°) - Inter-Rater	20.81 \pm 2.41	0.801	0.521–0.917	<0.001	1.08	2.11	11.6
MT - VL (cm) - Intra-Rater	2.11 \pm 0.31	0.987	0.969–0.995	<0.001	0.04	0.07	14.7
MT - VL (cm) - Inter-Rater	2.11 \pm 0.28	0.882	0.717–0.951	<0.001	0.09	0.19	13.1
MT - RF (cm) - Intra-Rater	1.89 \pm 0.28	0.848	0.633–0.937	<0.001	0.11	0.21	14.8
MT - RF (cm) - Inter-Rater	1.87 \pm 0.34	0.803	0.525–0.918	<0.001	0.15	0.29	17.9

FL: Fascicle Length; PA: Pennation Angle; MT: Muscle Thickness.

Table 3

Intra and inter-rater reliability results for the sit-to-stand (STS) test, area of force under the curves (AUC) in the first, third, and fifth repetitions of the STS, as well as average values for all STS test repetitions, and for the 6-min walk test (6MWT). Values of mean \pm standard deviation (SD), intraclass correlation coefficient (ICC), 95% confidence interval (95%CI), p-value, standard error of the measurement (SEM), minimum detectable change (MDC) and coefficient of variation (CV).

	Mean \pm SD	ICC	95%CI	p-value	SEM	MDC	CV (%)
STS (s) - Intra-Rater	8.92 \pm 0.87	0.950	0.879–0.979	<0.001	0.19	0.38	9.76
STS (s) - Inter-Rater	9.06 \pm 0.87	0.926	0.823–0.969	<0.001	0.24	0.46	9.61
STS - AUC Curve 1 (N·s) - Intra-Rater	113.7 \pm 24.6	0.898	0.754–0.958	<0.001	7.87	15.42	21.7
STS - AUC Curve 1 (N·s) - Inter-Rater	113.8 \pm 25.9	0.867	0.689–0.945	<0.001	9.45	18.53	22.8
STS - AUC Curve 3 (N·s) - Intra-Rater	113.7 \pm 24.7	0.885	0.723–0.952	<0.001	8.37	16.41	21.7
STS - AUC Curve 3 (N·s) - Inter-Rater	113.7 \pm 26.0	0.865	0.676–0.944	<0.001	9.55	18.72	22.9
STS - AUC Curve 5 (N·s) - Intra-Rater	116.8 \pm 30.7	0.669	0.202–0.862	0.007	17.64	34.58	26.3
STS - AUC Curve 5 (N·s) - Inter-Rater	116.7 \pm 30.8	0.586	0.003–0.828	0.025	19.83	38.87	26.4
STS - AUC Mean (N·s) - Intra-Rater	114.7 \pm 25.9	0.852	0.643–0.939	<0.001	9.96	19.53	22.6
STS - AUC Mean (N·s) - Inter-Rater	114.7 \pm 26.8	0.801	0.520–0.917	<0.001	11.95	23.42	23.4
6MWT (m) - Intra-Rater	691.2 \pm 65.7	0.977	0.945–0.990	<0.001	9.96	19.52	9.50
6MWT (m) - Inter-Rater	690.5 \pm 62.5	0.977	0.946–0.991	<0.001	9.49	18.59	9.06

Table 4

Inter-analyzer reliability results for the vastus lateralis (VL) and rectus femoris (RF) muscle architecture of the dominant inferior limb. Values of mean \pm standard deviation (SD), intraclass correlation coefficient (ICC), 95% confidence interval (95%CI), p-value, standard error of the measurement (SEM), minimum detectable change (MDC) and coefficient of variation (CV).

	Mean \pm SD	ICC	95%CI	p-value	SEM	MDC	CV (%)
FL - VL (cm) - Inter-Analyzer	6.51 \pm 0.83	0.769	0.445–0.904	0.001	0.40	0.78	12.7
PA - VL (°) - Inter-Analyzer	20.78 \pm 2.34	0.763	0.429–0.902	0.001	1.14	2.23	11.3
MT - VL (cm) - Inter-Analyzer	2.11 \pm 0.28	0.907	0.775–0.961	<0.001	0.09	0.17	13.4
MT - RF (cm) - Inter-Analyzer	1.89 \pm 0.31	0.990	0.975–0.996	<0.001	0.03	0.06	16.5

FL: Fascicle Length; PA: Pennation Angle; MT: Muscle Thickness.

to the instauration of fatigue at the lower limb muscles during the STS, modifying the motor gesture performance, which may be related to changes in the muscle strength capacity during STS (Roldan-Jimenez et al., 2015). In addition, caution is suggested in inferring possible measurement errors of the AUC during STS, since changes in muscle activation or motor control during the sit and stand task may affect the assessed STS force results (Roldan-Jimenez et al., 2015).

Regarding 6MWT results, the present study demonstrated that two evaluators with different previous experiences were able to obtain similar performance results for this functional test in healthy subjects. Probably the training procedures previous to data collection, as well as the use of the ATS recommendations for the 6MWT (Crapo et al., 2002), contributed to these findings. Our results also agree with previous studies performed with healthy subjects (Davi et al., 2014), as well as when compared to the intra-rater and inter-rater results from patients (Hansen et al., 2018).

5. Conclusion

In summary, the results of the present study demonstrated

strong intra-rater and inter-rater reliability for the knee extensor measures of force, torque and muscle activation. In addition, similar strong results were found for the functional STS and 6MWT tests. However, although most of VL and RF muscle architecture variables present strong reliability, more care is required to collect and analyze data from this technique.

6. Clinical relevance

- Neuromuscular outcomes and functional performance are regularly used to assess patients' improvement in clinical practice, and intra-rater and inter-rater reliability is of utter importance.
- The high intra-rater and inter-rater reliability obtained for the majority of the neuromuscular and functional outcomes evidences that the tests used in the present study can be used in different days by the same rater or in the same day by different rater.
- The moderate results in the between-raters reliability for some of the muscle architecture measures suggest that a higher training period should be observed when using this technique.

Credit Author Statement

FJL, GS and MAV contributed to the study conception and design; MAV and GS provided the material resources for the execution and were responsible for project coordination and institutional approval; FJL, FCS, IAP, LZO and DCSG conducted experiments; FJL, ESWN, DCSG and TM analyzed data. All authors contributed to the written manuscript, and read and approved the final manuscript version.

Funding

The authors received funding support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-Brazil) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil).

Declaration of competing interest

The authors declare that they have no conflict of interest. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

Acknowledgments

The authors declare no conflict of interest. FJL, IAP, ESWN and TM received a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-Brazil). GS and MAV received a research grant, not related to this study, from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil).

References

- Balshaw, T.G., Fry, A., Maden-Wilkinson, T.M., Kong, P.W., Folland, J.P., 2017. Reliability of quadriceps surface electromyography measurements is improved by two vs. single site recordings. *Eur. J. Appl. Physiol.* 117, 1085–1094. <https://doi.org/10.1007/s00421-017-3595-z>.
- Baroni, B.M., Geremia, J.M., Rodrigues, R., De Azevedo Franke, R., Karamanidis, K., Vaz, M.A., 2013. Muscle architecture adaptations to knee extensor eccentric training: rectus femoris vs. vastus lateralis. *Muscle Nerve* 48, 498–506. <https://doi.org/10.1002/mus.23785>.
- Bellet, R.N., Adams, L., Morris, N.R., 2012. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness—a systematic review. *Physiotherapy* 98, 277–286. <https://doi.org/10.1016/j.physio.2011.11.003>.
- Bohannon, R.W., 2011. Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults. *J. Strength Cond. Res.* 25, 3205–3207. <https://doi.org/10.1519/JSC.0b013e318234e59f>.
- Callaghan, M.J., McCarthy, C.J., Oldham, J.A., 2009. The reliability of surface electromyography to assess quadriceps fatigue during multi joint tasks in healthy and painful knees. *J. Electromyogr. Kinesiol.* 19, 172–180. <https://doi.org/10.1016/j.jelekin.2007.05.004>.
- Crapo, R.O., Casaburi, R., Coates, A.L., Enright, P.L., MacIntyre, N.R., McKay, R.T., Mottram, C., 2002. ATS statement: guidelines for the six-minute walk test. *Am. J. Respir. Crit. Care Med.* 166, 111–117. <https://doi.org/10.1164/ajrccm.166.1.at1102>.
- Davi, S.F., Arcuri, J.F., Labadessa, I.G., Pessoa, B.V., Costa, J.N.F., Sentanin, A.C., Di Lorenzo, V.A.P., 2014. Reprodutibilidade do teste de caminhada e do grau de 6 minutos em adultos jovens saudáveis. *Rev. Bras. Med. Esporte* 20, 214–2018. <https://doi.org/10.1590/1517-86922014200301714>.
- Eden, M.M., Tompkins, J., Verheijde, J.L., 2018. Reliability and a correlational analysis of the 6MWT, ten-meter walk test, thirty second sit to stand, and the linear analog scale of function in patients with head and neck cancer. *Physiother. Theory Pract.* 34, 202–211. <https://doi.org/10.1080/09593985.2017.1390803>.
- Fragala, M.S., Kenny, A.M., Kuchel, G.A., 2015. Muscle quality in aging: a multi-dimensional approach to muscle functioning with applications for treatment. *Sports Med.* 45, 641–658. <https://doi.org/10.1007/s40279-015-0305-z>.
- Goldberg, A., Chavis, M., Watkins, J., Wilson, T., 2012. The five-times-sit-to-stand test: validity, reliability and detectable change in older females. *Aging Clin. Exp. Res.* 24, 339–344. <https://doi.org/10.1007/BF03325265>.
- Gough, J.G., Hamrell, M., 2010. Standard operating procedures (SOPs): how to write them to be effective tools. *Ther. Innov. Regul. Sci.* 44, 463–468. <https://doi.org/10.1177/009286151004400410>.
- Guerra-Balic, M., Oviedo, G.R., Javierre, C., Fortuno, J., Barnet-Lopez, S., Nino, O., Alamo, J., Fernhall, B., 2015. Reliability and validity of the 6-min walk test in adults and seniors with intellectual disabilities. *Res. Dev. Disabil.* 47, 144–153. <https://doi.org/10.1016/j.ridd.2015.09.011>.
- Guilhem, G., Cornu, C., Maffiuletti, N.A., Guevel, A., 2013. Neuromuscular adaptations to isolaod versus isokinetic eccentric resistance training. *Med. Sci. Sports Exerc.* 45, 326–335. <https://doi.org/10.1249/MSS.0b013e31826e7066>.
- Hansen, H., Beyer, N., Frolich, A., Godtfredsen, N., Bieler, T., 2018. Intra- and inter-rater reproducibility of the 6-minute walk test and the 30-second sit-to-stand test in patients with severe and very severe COPD. *Int. J. Chronic Obstr. Pulm. Dis.* 13, 3447–3457. <https://doi.org/10.2147/COPD.S174248>.
- Hashemi Oskoue, A., Paulin, M.G., Carman, A.B., 2013. Intra-session and inter-day reliability of forearm surface EMG during varying hand grip forces. *J. Electromyogr. Kinesiol.* 23, 216–222. <https://doi.org/10.1016/j.jelekin.2012.08.011>.
- Ishida, H., Suehiro, T., Suzuki, K., Watanabe, S., 2018. Muscle thickness and echo intensity measurements of the rectus femoris muscle of healthy subjects: intra and interrater reliability of transducer tilt during ultrasound. *J. Bodyw. Mov. Ther.* 22, 657–660. <https://doi.org/10.1016/j.jbmt.2017.12.005>.
- Kollock Jr., R.O., Onate, J.A., Van Lunen, B., 2010. The reliability of portable fixed dynamometry during hip and knee strength assessments. *J. Athl. Train.* 45, 349–356. <https://doi.org/10.4085/1062-6050-45.4.349>.
- Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Kwah, L.K., Pinto, R.Z., Diong, J., Herbert, R.D., 2013. Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review. *J. Appl. Physiol.* 114, 761–769. <https://doi.org/10.1152/japplphysiol.01430.2011>, 1985.
- Lienhard, K., Laueremann, S.P., Schneider, D., Item-Glatthorn, J.F., Casartelli, N.C., Maffiuletti, N.A., 2013. Validity and reliability of isometric, isokinetic and isoinertial modalities for the assessment of quadriceps muscle strength in patients with total knee arthroplasty. *J. Electromyogr. Kinesiol.* 23, 1283–1288. <https://doi.org/10.1016/j.jelekin.2013.09.004>.
- Lima, K.M., da Matta, T.T., de Oliveira, L.F., 2012. Reliability of the rectus femoris muscle cross-sectional area measurements by ultrasonography. *Clin. Physiol. Funct. Imag.* 32, 221–226. <https://doi.org/10.1111/j.1475-097X.2011.01115.x>.
- Lima, K.M.M., Oliveira, L.F., 2013. Confiabilidade das medidas de arquitetura do músculo vasto lateral pela ultrassonografia. *Motriz* 19, 217–223. <https://doi.org/10.1590/S1980-65742013000100022>.
- Marzilger, R., Legerlotz, K., Panteli, C., Bohm, S., Arampatzis, A., 2018. Reliability of a semi-automated algorithm for the vastus lateralis muscle architecture measurement based on ultrasound images. *Eur. J. Appl. Physiol.* 118, 291–301. <https://doi.org/10.1007/s00421-017-3769-8>.
- McCarthy, C.J., Callaghan, M.J., Oldham, J.A., 2008. The reliability of isometric strength and fatigue measures in patients with knee osteoarthritis. *Man. Ther.* 13, 159–164. <https://doi.org/10.1016/j.math.2006.12.003>.
- Merletti, R., Botter, A., Troiano, A., Merlo, E., Minetto, M.A., 2009. Technology and instrumentation for detection and conditioning of the surface electromyographic signal: state of the art. *Clin. Biomech.* 24, 122–134. <https://doi.org/10.1016/j.clinbiomech.2008.08.006>.
- Mong, Y., Teo, T.W., Ng, S.S., 2010. 5-repetition sit-to-stand test in subjects with chronic stroke: reliability and validity. *Arch. Phys. Med. Rehabil.* 91, 407–413. <https://doi.org/10.1016/j.apmr.2009.10.030>.
- Paul, S.S., Canning, C.G., 2014. Five-repetition sit-to-stand. *J. Physiother.* 60, 168. <https://doi.org/10.1016/j.jphys.2014.06.002>.
- Place, N., Maffiuletti, N.A., Martin, A., Lepers, R., 2007. Assessment of the reliability of central and peripheral fatigue after sustained maximal voluntary contraction of the quadriceps muscle. *Muscle Nerve* 35, 486–495. <https://doi.org/10.1002/mus.20714>.
- Puthoff, M.L., Saskowski, D., 2013. Reliability and responsiveness of gait speed, five times sit to stand, and hand grip strength for patients in cardiac rehabilitation. *Cardiopulm. Phys. Ther. J.* 24, 31–37. <https://doi.org/10.1097/01823246-201324010-00005>.
- Raj, I.S., Bird, S.R., Shield, A.J., 2012. Reliability of ultrasonographic measurement of the architecture of the vastus lateralis and gastrocnemius medialis muscles in older adults. *Clin. Physiol. Funct. Imag.* 32, 65–70. <https://doi.org/10.1111/j.1475-097X.2011.01056.x>.
- Roldan-Jimenez, C., Bennett, P., Cuesta-Vargas, A.I., 2015. Muscular activity and fatigue in lower-limb and trunk muscles during different sit-to-stand tests. *PloS One* 10, e0141675. <https://doi.org/10.1371/journal.pone.0141675>.
- Ruschel, C., Hauptenthal, A., Jacomet, G.F., Fontana Hde, B., Santos, D.P., Scoz, R.D., Roesler, H., 2015. Validity and reliability of an instrumented leg-extension machine for measuring isometric muscle strength of the knee extensors. *J. Sport Rehabil.* 24, 1–4. <https://doi.org/10.1123/jsr.2013-0122>.
- Santos, A.N., Pavao, S.L., Avila, M.A., Salvini, T.F., Rocha, N.A.C.F., 2013. Reliability of isokinetic evaluation in passive mode for knee flexors and extensors in healthy children. *Braz. J. Phys. Ther.* 17, 112–120. <https://doi.org/10.1590/S1413-35552012005000074>.
- Santos, R., Armada-da-Silva, P.A.S., 2017. Reproducibility of ultrasound-derived muscle thickness and echo-intensity for the entire quadriceps femoris muscle. *Radiography* 23, e51–e61. <https://doi.org/10.1016/j.radi.2017.03.011>.
- Schurr, K., Sherrington, C., Wallbank, G., Pamphlett, P., Olivetti, L., 2012. The minimum sit-to-stand height test: reliability, responsiveness and relationship to leg muscle strength. *Clin. Rehabil.* 26, 656–663. <https://doi.org/10.1177/0269215511427323>.
- Schwenk, M., Gogulla, S., Englert, S., Czempik, A., Hauer, K., 2012. Test-retest reliability and minimal detectable change of repeated sit-to-stand analysis using

- one body fixed sensor in geriatric patients. *Physiol. Meas.* 33, 1931–1946. <https://doi.org/10.1088/0967-3334/33/11/1931>.
- Silva, C.R.S., Costa, A.D.S., Rocha, T., Lima, D.A.M., Nascimento, T.D., Moraes, S.R.A., 2018. Quadriceps muscle architecture ultrasonography of individuals with type 2 diabetes: reliability and applicability. *PloS One* 13, e0205724. <https://doi.org/10.1371/journal.pone.0205724>.
- Sorbie, G.G., Williams, M.J., Boyle, D.W., Gray, A., Brouner, J., Gibson, N., Baker, J.S., Easton, C., Ugbole, U.C., 2018. Intra-session and inter-day reliability of the myon 320 electromyography system during sub-maximal contractions. *Front. Physiol.* 9, 309. <https://doi.org/10.3389/fphys.2018.00309>.
- Suzuki, T., 2015. Reliability of measurements of knee extensor muscle strength using a pull-type hand-held dynamometer. *J. Phys. Ther. Sci.* 27, 967–971. <https://doi.org/10.1589/jpts.27.967>.
- Toonstra, J., Mattacola, C.G., 2013. Test-retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. *J. Sport Rehabil.* 22, 1–5. <https://doi.org/10.1123/jsr.2013.TR7>.
- Tveter, A.T., Dagfinrud, H., Moseng, T., Holm, I., 2014. Measuring health-related physical fitness in physiotherapy practice: reliability, validity, and feasibility of clinical field tests and a patient-reported measure. *J. Orthop. Sports Phys. Ther.* 44, 206–216. <https://doi.org/10.2519/jospt.2014.5042>.
- Wang, T.H., Liao, H.F., Peng, Y.C., 2012. Reliability and validity of the five-repetition sit-to-stand test for children with cerebral palsy. *Clin. Rehabil.* 26, 664–671. <https://doi.org/10.1177/0269215511426889>.
- Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Condit. Res.* 19, 231–240. <https://doi.org/10.1519/15184.1>.
- Winter, D.A., 2009. *Biomechanics and Motor Control of Human Movement*, fourth ed. John Wiley & Sons. <https://doi.org/10.1002/9780470549148>.