

TEST-RETEST RELIABILITY OF ELECTROMYOGRAPHIC SIGNAL PARAMETERS USED TO EVALUATE NEUROMUSCULAR FATIGUE IN QUADRICEPS FEMORIS MUSCLE

Cristiano Rocha da Silva^{1,2}, Danilo de Oliveira Silva¹, Ronaldo Valdir Briani¹, Marcella Ferraz Pazzinato¹, Deisi Ferrari^{1,3}, and Fábio Mícolis de Azevedo¹

¹Laboratory of Biomechanics and Motor Control – FCT/UNESP, Presidente Prudente, Brazil

²Neuroscience Program and Biomedical Engineering Laboratory, Univ. of São Paulo, EPUSP, PTC, São Paulo, Brazil.

³University of São Paulo, Post-graduation Program Interunits Bioengineering, EESC/FMRP/IQSC-USP, São Carlos, Brazil

Original scientific paper

UDC: 612:796.012:519.2

Abstract:

The purpose of this study was to analyze the test-retest reliability of the median frequency (MDF) and root mean square (RMS) used to determine the onset of neuromuscular fatigue (NF) during sustained fatiguing contractions of the quadriceps. Eighteen healthy men were tested on two days, and electromyographic (EMG) signals were recorded from the vastus medialis (VM), rectus femoris (RF) and vastus lateralis (VL) during sustained isometric contractions at 20 and 70% of maximum voluntary contractions (MVC) held until exhaustion. The reliability of endurance time was excellent at 20% MVC and poor at 70% MVC. EMG variables were evaluated: (1) at the beginning of the test; (2) at NF; and (3) at the end of the test. The NF time values presented poor reliability. The MDF has shown, in general, poor reliability at 20 and 70% MVC, whereas the RMS reliability presented better results for both loads, especially for RF, followed by the VM and VL muscles. The MDF and RMS values extracted from NF showed poor reliability at 20 and 70% MVC, which suggests caution in using these variables extracted from the EMG signal to determine the onset of NF.

Key words: *electromyography, isometric contraction, reliability, reproducibility of results, quadriceps muscle*

Introduction

Neuromuscular fatigue (NF) is defined as the decline in muscle performance during and after repetitive and sustained contractions (Babault, Desbrosses, Fabre, Michaut, & Pousson, 2006). A widely accepted method to assess NF is surface electromyography (Rogers & MacIsaac, 2013). Merletti, Lo Conte, and Orizio (1991) report that fatigue may be interpreted as an attribute of a specific contraction and can therefore be represented by a number or index associated with the contraction. Thus, the variables related to the amplitude of electromyographic (EMG) signal (e.g. root mean square [RMS]) and its frequency content (e.g. median frequency of the power spectrum [MDF]) are commonly used to assess physiologically relevant aspects such as muscle-activation level and development of NF (Briani, et al., 2015; Ferrari,

Kuriki, Silva, Alves, & Mícolis de Azevedo, 2014; Contessa, Adam, & De Luca, 2009; Garcia, et al., 2010).

There is a demand for studies related to the application of techniques investigating the moment of transition between muscular work without NF and the onset of NF (De Luca, 1984, 1997). Several studies have evaluated the RMS and MDF values only at the beginning and end of the fatigue tests (Callaghan, McCarthy, & Oldham, 2009; Hedayatpour, Arendt-Nielsen, & Farina, 2008; Mathur, Eng, & MacIntyre, 2005; Santos, et al., 2008; Watanabe & Akima, 2010). However, the concept of NF onset used by the authors cited is based on the point at which a contraction can no longer be maintained (the failure point). Many studies use force output of a muscle as the index of NF. The use of the failure point brings some practical disadvantages,

for example, fatigue is detected only after it has occurred (De Luca, 1984, 1997). In a recent study (Silva, et al., 2014b) we evaluated the onset of NF through MDF and RMS behavior during sustained fatiguing contractions of the quadriceps, i.e. not the decrease of the output force. This analysis demonstrated that it is possible to find significant alterations in the EMG signal prior to the reduction of output force. However, this being a new analysis, we believe it is important and necessary to assess reliability of these variables. Validation of the technique is necessary to assure that data recorded in different laboratories and by different operators are comparable in terms of reliability (Rainoldi, Bullock-Saxton, Cavarretta, & Hogan, 2001).

Several works regarding the reliability of EMG frequency and amplitude parameters in sustained submaximal contractions of the quadriceps femoris muscle can be found in the literature of the last decade (Callaghan, et al., 2009; Kollmitzer, Ebenbichler, & Kopf, 1999; Mathur, et al., 2005; Rainoldi, et al., 2001). No published studies were found in the databases consulted that verified the reliability of the EMG signal between the beginning and end of the fatigue tests, i.e. identifying whether the reliability at the onset of NF is different from the reliability found at the beginning and end of the test. Importantly, many neurophysiological mechanisms are perturbed before the body feels the effects of fatigue, and these changes sometimes constitute an early warning of fatigue (Boyas & Guevel, 2011), which can be verified by decrements in the force generated during testing (Hedayatpour, et al., 2008; Mathur, et al., 2005). NF then develops progressively until the muscle is no longer able to perform the requested task (Boyas & Guevel, 2011). It is important to observe that the reduction in force is first observed only at the end of testing, at which time the muscle is already working under fatigue.

Thus, the purpose of this study was to evaluate the reliability of variables obtained in the time and frequency domain of EMG signals obtained at three different points of the fatigue tests: (1) at the beginning of the test; (2) at the onset point of NF; and (3) at the end of the test.

Methods

Participants

The number of participants needed for the reliability study was based on calculations described in a previous study (Callaghan, et al., 2009). For test sessions the minimum number of participants needed was 14. This study evaluated 18 male volunteers (age 21 ± 2 years, body mass 79.1 ± 13 kg, body height 166.4 ± 6.5 cm, body mass index – BMI 25.3 ± 3.3 kg/m²) who were recreationally active, with no previous history of trauma, surgery, pain, or any other neuromuscular knee disorder. The domi-

nant limb was determined based on each subject's preference while performing an act of kicking (Silva, et al., 2014b). This study was approved by the local University Ethics Committee and it conformed to the latest amendments set by the Declaration of Helsinki. All subjects provided a written informed consent prior to the tests.

Experimental procedure

The subjects were seated on a knee extensor table (VITTALY[®], model convergent[®], São José do Rio Preto – SP, Brazil) with the back adjustments and support. The knee-joint position was maintained at 60° of flexion, and the hip at 90° of flexion. The positions of the hip and knee were confirmed using a universal goniometer. The trunk, pelvis and knee were firmly strapped to the table with a seat-belt (Silva, et al., 2014b).

Three maximum voluntary isometric knee-extension contractions (MVC) were performed, and verbal encouragement was provided for a six-second period to induce the subject to reach their highest level in each trial. Visual feedback of the produced force was also provided. A five-minute break was given between MVC trials (Rainoldi, Falla, Mellor, Bennell, & Hodges, 2008). The first and the last two second of the MVC signal were ignored. The final MVC force was determined by the overall average value of the peak force recorded over the three attempts. This value was used to calculate loads of the fatigue tests.

Two days after the MVC test, the subjects were asked to perform a submaximal knee-extension contraction at 20 and 70% MVC held until exhaustion. The choice of the testing order was random and the exhaustion moment was considered to be the point at which the subject completely ceased exercise, or when the force output deviated more than 10% from the target load (Callaghan, et al., 2009; Silva, et al., 2014a, 2014b). The order of the fatigue tests was different between the tests days. To assist in the control of load intensity, the force applied was monitored in real time by a load cell, representing a visual display of feedback of the force output. A thirty-minute rest period was provided between the submaximal contractions. The fatigue tests were performed twice with a two-day interval between them.

EMG signal measurement

The EMG signals were recorded from three pairs of surface electrodes in bipolar configuration (gain = 1000, sampling frequency = 4000Hz), on the skin covering the quadriceps femoris muscle (Medi-trace[®] Ag-AgCL, Mansfield, MA, 10 mm radius disc electrodes, fixed interelectrode distance of 20 mm) by a signal conditioner module (BIO EMG 1000 model, LYNX[®], Electronic Technology Ltda, Sao Paulo – SP, Brazil). The electrode cable circuit

had a preamplifier with a gain of 20 times, Common Mode Rejection greater than 80 dB and impedance of $1\text{K}\Omega$. The areas chosen for electrode placement (electrical stimulation technique of motor points) were prepared by shaving, abraded by fine sandpaper, and cleaned with isopropyl alcohol; electrodes were positioned parallel to the vastus medialis (VM), rectus femoris (RF) and vastus lateralis (VL) muscles and fixed approximately 2 cm from the motor point in the direction of the muscle fibers (De Luca, 1997). During the fatigue test, the force applied by the subject was monitored using a load cell (model MM[®], KRATOS[®], Cotia – SP, Brazil) synchronized with the EMG signals. A specific channel of the module signal conditioner was set to acquire the signals from the load cell with a second-order Butterworth low-pass filter with a cutoff frequency of 100Hz.

The fixation of electrodes at the same location is an important condition for reducing inherent variability in measurement due to their position, especially in studies of reproducibility. In order to ensure the same position of the electrodes over the three days of data collection, we developed a feedback system which consisted of a sheet of acetate, on which the upper border of the patella, the shaved area of each muscle portions and the motor point were marked. When the subject returned for the next test day, the acetate sheet was aligned with the upper border of the patella and the shaved areas, and only then the points were marked (Silva, et al., 2012).

Signal processing

The off-line processing of the EMG signals was performed by MATLAB (The MathWorks[®] Inc., Natick, MA). The EMG data were bandpass filtered using a fourth order, zero lag, Butterworth filter with 20-500 Hz cutoff frequencies. The power spectral density and frequency characteristics were determined using Welch's averaged periodogram method and the fast Fourier transform technique. Following this process, the MDF and RMS shifts were computed. To indicate the presence of neuromuscular fatigue, it is important to determine how the final MDF and RMS values relate to the initial MDF and RMS values. During fatigue protocols an increase in RMS values and a decrease in MDF values (Mathur, et al., 2005; Silva, et al., 2014b) are observed. Thus, the MDF was normalized by averages obtained from five initial values of the signal, and the RMS was normalized by the RMS obtained in the MVC tests.

For each test session, the following epochs of the EMG signal were selected as described in Silva et al. (2014b): (1) beginning of the test; (2) at NF onset; and (3) end of the test. At the beginning of the test, the initial MDF and RMS were calculated considering the average of the first five samples of

MDF and RMS from the EMG signal. The final MDF and final RMS were calculated considering the average of the final five samples of MDF and RMS from the EMG signal. For the NF measure, two methodologies were used in this study: (1) the analysis of the MDF slope coefficient (NF_{MDF}); and (2) analysis of the RMS slope coefficient (NF_{RMS}). The steps taken for this analysis were: (1) a linear fit was applied for every five samples of MDF and RMS data during the fatigue exercise, progressively, and always from the origin until exhaustion, (2) on each muscle the coefficient slope of the adjusted MDF and RMS values, as well as the value for the *t*-test and their significance ($p < .05$) were analyzed; and (3) NF was established as the first section where the slope coefficient showed a significant difference from zero. MDF was calculated in NF_{MDF} and RMS was calculated in NF_{RMS} (Silva, et al., 2014b).

Statistical analysis

Statistical analysis was performed using the Statistical Package for Social Sciences software program for Windows (version 18.0, SPSS, INC., Chicago, IL). The test-retest relative reliability of the measurements was evaluated using the intraclass correlation coefficient (ICC) with 95% confidence intervals (95% CI) (De Oliveira Silva, et al., 2015a). This study used the ICC_{2,1} model (two way random with absolute agreement), which in the current practice considers ICC values in the range of .80 to 1.00 as 'excellent reliability' and ICC values in the range of .60 to .80 as 'good reliability', while an ICC less than .60 reflects 'poor reliability' (Al-Zahrani, et al., 2009; Callaghan, et al., 2009). The *F*-test values for the between-subjects effect of reliability analysis were also verified, which was not reported in the original analysis of variance (ANOVA) table (Portney & Watkins, 2009). In addition to ICC, the standard error of measurement (SEM) and the minimum difference (MD) to be considered 'real' were calculated. The SEM was used to express absolute reliability of the measure and was calculated as the square root of the mean square error term from the ANOVA, and had the same unit as the tested variable. The MD was defined as a clinical applicability index, expressed as the same units of the measurement ($MD = SEM \times 1.96$). The SEM ($SEM / \text{grand mean} \times 100$) and the MD ($MD / \text{grand mean} \times 100$) were expressed as a percentage of the mean value for the measure to allow for comparison of absolute reliability between measurements (Al-Zahrani, et al., 2009; Mathur, et al., 2005; Weir, 2005).

Results

The reliability of endurance time during fatigue tests was excellent at 20% MVC ($ICC_{2,1} = .83$, 95% CI = .60-.93, $p[F\text{-test}] < .001$) and poor at 70% MVC

Table 1. Reliability data: Group ICC, SEM and MD for the normalized time (%) of NF_{MDF} and NF_{RMS} obtained in fatigue tests at 20 and 70% MVC for all muscles

	NF_{MDF}		NF_{RMS}	
	20% MVC	70% MVC	20% MVC	70% MVC
VM				
ICC _{2,1} (95% CI)	-.14 (-.58 – .37)	.49 (.05 – .77)*	.57 (.16 – .82)*	.70 (.32 – .89)*
SEM%	73.40	39.27	45.95	35.53
MD%	203.27	108.74	127.24	98.39
RF				
ICC _{2,1} (95% CI)	-.25 (-.69 – .27)	.39 (-.07 – .72)*	.08 (-.37 – .51)	.56 (-.04 – .84)*
SEM%	84.14	50.33	69.43	45.92
MD%	232.98	139.36	192.26	127.20
VL				
ICC _{2,1} (95% CI)	.32 (-.13 – .68)	.29 (-.12 – .65)	.65 (.27 – .86)*	.42 (-.05 – .74)*
SEM%	53.18	40.29	45.61	45.97
MD%	147.34	122.25	126.33	127.31

Abbreviations: NF_{MDF} – onset of muscle fatigue obtained by median frequency behavior; NF_{RMS} – onset of muscle fatigue obtained by root mean square behavior; MVC – maximum voluntary contraction; VM – vastus medialis; RF – rectus femoris; VL – vastus lateralis; ICC – intraclass correlation coefficient; 95% CI – 95% confidence intervals, SEM% – standard error of measurement (as percentage of grand mean), MD% – smallest detectable difference (as percentage of grand mean).

* ICC with significant values for F-test, $p < .05$.

Table 2. Reliability data: Group ICC, SEM and MD for MDF and RMS values obtained in fatigue tests at 20 and 70% MVC for all muscles

	VM			RF			VL		
	ICC _{2,1} (95% CI)	SEM%	MD%	ICC _{2,1} (95% CI)	SEM%	MD%	ICC _{2,1} (95% CI)	SEM%	MD%
MDF									
	00			00					
20% MVC									
I_{MDF}	-.12 (-.61 – .40)	5.40	14.81	-.35 (.68 – .13)	3.98	11.02	-.18 (-.66 – .34)	5.69	15.74
NF_{MDF}	.13 (-.41 – .12)	5.99	16.59	.10 (-.37 – .54)	4.89	13.55	.08 (.45 – .55)	5.86	16.23
F_{MDF}	.69 (.30 – .88)*	6.65	18.40	.55 (.10 – .81)*	5.51	15.27	.42 (-.08 – .75)*	8.81	24.39
70% MVC									
I_{MDF}	.09 (-.40 – .53)	4.95	13.72	.38 (-.09 – .71)	3.51	9.74	-.14 (-.60 – .36)	3.86	10.70
NF_{MDF}	.23 (-.24 – .61)	4.72	14.69	.30 (-.19 – .67)	5.08	14.08	.26 (-.15 – .62)	4.87	13.50
F_{MDF}	.56 (.14 – .81)*	6.80	19.03	.49 (.03 – .78) *	8.82	24.41	.48 (.02 – .77)*	6.71	18.57
RMS									
20% MVC									
I_{RMS}	.71 (.34 – .89)*	16.57	45.94	.76 (.45 – .90)*	16.03	44.44	.74 (.39 – .90)*	19.71	54.55
NF_{RMS}	.68 (.30 – .88)*	17.99	49.81	.76 (.46 – .91)*	16.64	46.04	.56 (.13 – .82)*	22.57	62.66
F_{RMS}	.89 (.71 – .96)*	6.21	12.82	.87 (.67 – .95)*	16.22	44.92	.84 (.59 – .94)*	20.53	56.89
70% MVC									
I_{RMS}	.41 (-.07 – .73)*	12.71	17.53	.83 (.60 – .93)*	13.89	38.46	.59 (.17 – .82)*	17.84	49.44
NF_{RMS}	.44 (-.02 – .75) *	17.63	48.84	.87 (.71 – .95)*	13.10	36.30	.40 (-.08 – .72)*	19.55	54.14
F_{RMS}	.64 (.26 – .85)*	16.81	45.56	.76 (.48 – .90)*	19.16	53.05	.58 (.16 – .82)*	21.68	60.04

Abbreviations: VM – vastus medialis; RF – rectus femoris; VL – vastus lateralis; ICC – intraclass correlation coefficient; 95% CI – 95% confidence intervals; SEM% - standard error of measurement (as percentage of grand mean); MD% – smallest detectable difference (as percentage of grand mean); MDF – median frequency; MVC – maximum voluntary contraction; I_{MDF} – initial MDF values; NF_{MDF} – MDF values extracted from the NF obtained by median frequency behavior; F_{MDF} – final MDF values; RMS – root mean square; I_{RMS} – initial RMS values; NF_{RMS} – RMS values extracted from the NF obtained by root mean square behavior; F_{RMS} – final RMS values.

* ICC with significant values for F-test, $p < .05$.

($ICC_{2,1}=.57$, 95% CI=.15-.81, $p[F\text{-test}]=.006$). The endurance time showed moderate SEM and MD values (20% MVC: SEM=12.51%, MD=34.67%; 70% MVC: SEM=12.45%, MD=34.49%), indicating a moderate measurement error and reproducibility only at the low load.

Table 1 shows the test-retest reliability of NF_{MDF} and NF_{RMS} time values (normalized by absolute endurance time [%]) for VM, RF and VL muscles at 20 and 70% MVC in the two days. All analyses showed, in general, low ICC values and high SEM and MD values indicating poor reliability across the days, in which better NF_{RMS} results were observed compared with NF_{MDF} for all muscles and both loads (Table 1).

Table 2 demonstrates poor reliability for the initial MDF at 20 and 70% MVC for the three muscles; however, the SEM (ranging from 3 to 5%) and MD values (ranging from 9 to 15%) were low. The same occurred with the NF_{MDF} showing poor reliability with low SEM (ranging from 4 to 5%) and MD (ranging from 13 to 16%) values for 20 and 70% MVC. The final MDF values also showed, in general, poor reliability, with low values of SEM and MD (Table 2).

The reliability of the RMS at 20% MVC showed good to excellent reliability values accompanied by low values of SEM and MD for the three muscles (except for the NF_{RMS} from the VL). The ICC values at 70% MVC were generally poor for the VM and VL and good to excellent for the RF muscle (ranging from 0.76 to 0.87). Regarding the SEM and MD values, the three muscles showed low values for the 20% MVC tests and high values for the 70% MVC tests (Table 2).

Discussion

The endurance time values showed different values of reliability and were revealed to be dependent on the load applied during the test. For the fatigue tests performed at 70% MVC the time values showed poor reliability and at 20% MVC excellent reliability. Mathur et al. (2005) found discordant reliability values in fatigue tests at 80% MVC ($ICC=.85$), but in sustained contractions at 20% MVC the reliability was similar ($ICC=.96$) demonstrating much larger time values (379±138 seconds for the first day; 400±198 seconds for the second day). It is suggested that the explanation for the time values in the current study being lower in both loads and accompanied by a smaller variability is related to the selection criteria for the interruption of the fatigue test. In the present study the exhaustion moment was considered to be the point at which the force output appeared to deviate more than 10% from the target load, whereas Mathur et al. (2005) considered a decrease exceeding 20% of the target load, thus enabling individuals to execute fatigue tests for a much longer period of time.

The time values found in the NF_{MDF} and NF_{RMS} showed poor reliability for all muscles and both loads. These values had high relative and absolute variability; however, this is one of the first studies to address the reliability of the onset of NF by means of RMS and MDF behavior since similar studies were not found in the databases consulted to discuss such behaviors.

The initial and final MDF values at both loads demonstrated poor reliability for the three muscles. The data is not in accordance with Mathur et al. (2005), who found moderate to high reliability for the initial and final MDF values at low and high contraction intensities. Other studies using protocols with sustained submaximal contractions showed good reliability for the initial MDF values in extensor muscles of the trunk (Elfving, Németh, Arvidsson, & Lamontagne, 1999), elbow extensors (Bilodeau, Arsenault, Gravel, & Bourbonnais, 1994) and quadriceps femoris muscles (Rainoldi, et al., 2001). Kollmitzer et al. (1999) found greater MDF reliability of the RF at 50% MVC than the VM and VL. Mathur et al. (2005) also showed that the RF muscle had the largest relative and absolute reliability for the initial and final values of MDF compared with the VL and VM at 20 and 80% MVC. The authors of these two studies reported that the thigh position during fatigue testing can account for the increased variability in EMG signals of the VL and VM during fatigue tests. A slight internal or external hip rotation can change the way each of these muscles is recruited, increasing the variability of data between days (Mannion & Dolan, 1996).

For RMS reliability values, in general, all measures showed good to high reliability at 20% MVC, primarily for the RF. For 70% MVC, the RF muscles presented good to high reliability, while the VM and VL showed, in general, poor reliability. These findings disagree with Mathur et al. (2005), who found good reliability for initial and final RMS at both loads, and agree with the same authors regarding better performance at 20% MVC compared with 70% MVC. As the amplitude of the EMG signal is related to force, this may reflect the differences in the ability to hold the force steady at low *versus* high forces. An inverse relationship exists between force production and force variability (i.e. steadiness of force); with higher forces showing greater force variations and hence, an increase in within-subject variability. This may have contributed to a greater degree of variability in the amplitude of EMG between-days and accounted for the lower ICCs for the high intensity fatigue tests (Mathur, et al., 2005).

In general, the NF_{MDF} and NF_{RMS} reliability values found were poor at 20 and 70% MVC with moderate-to-high SEM and MD values. It is understood that low-moderate-high ICC values accompanied by low absolute reliability are an indication that

such measures are not adequate for future clinical application (Rainoldi, et al., 2001).

Reliability considerations

In the present study, the ICC was used to verify the reliability of variables extracted from the EMG signal during sustained isometric contractions. It is possible to note in some tables the presence of negative ICC values and significant values for the Fisher test (F -test). It is known that performance of the ICC is dependent on the variability present between the variables obtained from each subject, being an index of the relative reliability of the measurements, expressing the ratio of between-subject variance to within-subject variance (Mathur, et al., 2005). One reason for low ICC values can be explained by the fact that the variability among subjects' scores must be large to demonstrate reliability (De Oliveira Silva, et al., 2015b). A lack of variability can occur when samples are homogeneous. This effect can be verified by looking at the significance of the between-subjects variance in the analysis of covariance. If subjects' scores are homogeneous, this source of variance will not be significant. It has been shown that when the between-subjects variance is not significant, the actual limits of the ICC do not match the theoretical limits of .00 and 1.00. In fact, it is possible for ratios to range from negative to positive infinity. When a negative ICC is obtained, the value cannot be considered valid. Therefore, it is imperative that researchers are aware of the extent to which scores will naturally vary, and try to obtain heterogeneous samples whenever possible (Portney & Watkins, 2009). The test which assesses the significance of variance between subjects is the F -test. The F value in ICC analysis is commonly obtained by dividing the between-subjects mean square by the error mean square. When the result of this test is significant ($p < .05$), it tells us that the subjects are different from each other. This is a necessary condition for reliability testing, because the validity of the ICC will be suspect if the F -test is not significant (Portney & Watkins, 2009).

Thus, based on the literature, in the present study the F -test was used in all reliability analyses. In general, no differences were presented between individuals for the NF_{MDF} with negative and poor values of ICC. For the NF_{RMS} the reliability was better, but nonetheless presented poor reliability values (Table 1). For the MDF and RMS reliability analysis (Table 2), significant F -test values were noted for all the analyses performed on RMS data. In the same table, the MDF showed few significant F -test values, with some negative ICC values.

A possible reason why the ICCs were better for RMS than MDF is because the magnitude of an ICC reflects the between-subjects variability. It is known that greater between-subjects variability will result in a higher ICC when compared to lower

between-subjects variability. For this reason, the ICC is considered to be context-specific for a test (Herda, et al., 2008; Weir, 2005; De Oliveira Silva, et al., 2015c). Rainoldi et al. (2001) reported the same problems found in their study. The authors reported that, if the within-subject variability is comparable with the between-subject variability, the ICC values may become negative. Although this apparent nonsense is clearly explained from a mathematical point of view, it limits the use of the ICC, especially for clinical applications. Given this limitation, the authors used other statistical techniques to assess the reliability of measurements, performing the F -test for the ratio between the mean squared error due to the subject's differences and the sum of mean squared errors due to trial and day variations. The authors demonstrated that scientific community is interested in developing other indices of reliability, not to replace the ICC, but rather to supplement the analysis when the calculation of the ICC is not possible.

However, it is also important to emphasize that the ICC is an index that does not provide information to assess real changes in the variable investigated, i.e. beyond the limitations that may appear depending on the behavior of the measurements' variability, it is configured as an index difficult to be clinically interpreted. Therefore, it has been suggested by many authors that the absolute error measurements such as SEM and MD should be used along with the ICC, thus allowing better identification of the size of errors in the size of the change in the variable under study (Al-Zahrani, et al., 2009; Callaghan, et al., 2009; Mathur, et al., 2005; Weir, 2005).

Although this study found relevant results regarding the reliability of the EMG signal during fatigue tests, several limitations need to be taken into account when considering these findings. Firstly, there are inherent issues with bipolar surface electrodes, such as cross-talk. Even using the electrical stimulation technique and an electrode feedback system, these methods are not able to completely eliminate the cross-talk artifact between the VM, RF and VL muscles. Secondly, there were differences in the execution strategy between subjects. The knee joint is a complex articular structure controlled by a number of interacting elements including muscles, tendons and ligaments. Subjects could produce similar torque output by using subtly different limb positions and muscle activation patterns. We minimized this influence by standardizing the test position for all subjects by fixing the trunk, pelvis, and knee joint. Thirdly, the rest period between fatigue tests performed on the same day (separated by 30 minutes) may have influenced the performance of the second test. Lastly, the two-day interval between fatigue tests was within the 72-hour window in which one would expect

delayed onset of muscle soreness, which could alter the endurance time of the tests.

Although the RMS reliability demonstrated better results compared with the MDF reliability, EMG power spectrum variables are a very important way to assess fatigue conditions. Future studies should evaluate other variables obtained in the frequency domain of the EMG signal that address the same goal, always checking if the reproducibility of the measures remains reliable. Additionally, further studies should investigate the influence of the muscle fiber type proportions and different isometric fatigue protocols on the time and frequency domains of the EMG signal, as well as examine whether reliability remains suitable during a fatigue test.

The parameters extracted from the EMG signal during sustained submaximal isometric contractions demonstrated different values of reliability. The results showed poor reliability for NF determined through RMS and MDF behavior in this study. In general, MDF presented poor reliability values during the fatigue tests, whereas RMS showed better reliability considering the proposed experimental design, especially for RF, followed by the VM and VL muscles. Furthermore, surface electromyography should continue to be used to assess changes in fatigability of muscles; however, the variables need to be chosen carefully in order to ensure that such measures can be used properly for future studies or clinical applications.

References

- Al-Zahrani, E., Gunasekaran, C., Callaghan, M., Gaydecki, P., Benitez, D., & Oldham, J. (2009). Within-day and between-days reliability of quadriceps isometric muscle fatigue using mechanomyography on healthy subjects. *Journal of Electromyography and Kinesiology*, 19(4), 695-703.
- Babault, N., Desbrosses, K., Fabre, M.-S., Michaut, A., & Pousson, M. (2006). Neuromuscular fatigue development during maximal concentric and isometric knee extensions. *Journal of Applied Physiology*, 100(3), 780-785.
- Bilodeau, M., Arsenault, A.B., Gravel, D., & Bourbonnais, D. (1994). EMG power spectrum of elbow extensors: A reliability study. *Electromyography and Clinical Neurophysiology*, 34(3), 149-158.
- Boyas, S., & Guevel, A. (2011). Neuromuscular fatigue in healthy muscle: Underlying factors and adaptation mechanisms. *Annals of Physical and Rehabilitation Medicine*, 54(2), 88-108.
- Briani, R.V., De Oliveira Silva, D., Pazzinatto, M.F., Albuquerque, C.E., Ferrari, D., Aragao, F.A., & De Azevedo, F.M. (2015). Comparison of frequency and time domain electromyography parameters in women with patellofemoral pain. *Clinical Biomechanics (Bristol, Avon)*, 30(3), 302-307.
- 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. *Journal of Electromyography and Kinesiology*, 19(1), 172-180.
- Contessa, P., Adam, A., & De Luca, C.J. (2009). Motor unit control and force fluctuation during fatigue. *Journal of Applied Physiology*, 107(1), 235-243.
- De Luca, C.J. (1984). Myoelectrical manifestations of localized muscular fatigue in humans. *Critical Reviews in Biomedical Engineering*, 11(4), 251-279.
- De Luca, C.J. (1997). The use of surface electromyography in biomechanics. *Journal of Applied Biomechanics*, 13(2), 135-163.
- De Oliveira Silva, D., Briani, R.V., Pazzinatto, M.F., Ferrari, D., Aragão, F.A., & De Azevedo, F.M. (2015a). Reduced knee flexion is a possible cause of increased loading rates in individuals with patellofemoral pain. *Clinical Biomechanics*, 30(9), 971-975.
- De Oliveira Silva, D., Briani, R.V., Pazzinatto, M.F., Ferrari, D., Aragão, F.A., Albuquerque, C.E., Alves, N., & Azevedo, F.M. (2015c). Reliability and differentiation capability of dynamic and static kinematic measurements of rearfoot eversion in patellofemoral pain. *Clinical Biomechanics*, 30(2), 144-148.
- De Oliveira Silva, D., Briani, R.V., Pazzinatto, M.F., Gonçalves, A.V., Ferrari, D., Aragão, F.A., & De Azevedo, F.M. (2015b). Q-angle static or dynamic measurements, which is the best choice for patellofemoral pain? *Clinical Biomechanics*, 30(10), 1083-1087.
- Elfvig, B., Németh, G., Arvidsson, I., & Lamontagne, M. (1999). Reliability of EMG spectral parameters in repeated measurements of back muscle fatigue. *Journal of Electromyography and Kinesiology*, 9(4), 235-243.
- Ferrari, D., Kuriki, H.U., Silva, C.R., Alves, N., & Micolis de Azevedo, F. (2014). Diagnostic accuracy of the electromyography parameters associated with anterior knee pain in the diagnosis of patellofemoral pain syndrome. *Archives of Physical Medicine and Rehabilitation*, 95(8), 1521-1526.
- Garcia, F.R., Azevedo, F.M., Alves, N., Carvalho, A.C., Padovani, C.R., & Negrao Filho, R. F. (2010). Effects of electrical stimulation of vastus medialis obliquus muscle in patients with patellofemoral pain syndrome: An electromyographic analysis. *Revista Brasileira de Fisioterapia*, 14(6), 477-482.

- Hedayatpour, N., Arendt-Nielsen, L., & Farina, D. (2008). Non-uniform electromyographic activity during fatigue and recovery of the vastus medialis and lateralis fatigue and recovery muscles. *Journal of Electromyography and Kinesiology*, 18(3), 390-396.
- Herda, T.J., Ryan, E.D., Beck, T.W., Costa, P.B., DeFreitas, J.M., Stout, J.R., et al. (2008). Reliability of mechanomyographic amplitude and mean power frequency during isometric step and ramp muscle actions. *Journal of Neuroscience Methods*, 171(1), 104-109.
- Kollmitzer, J., Ebenbichler, G.R., & Kopf, A. (1999). Reliability of surface electromyographic measurements. *Clinical Neurophysiology*, 110(4), 725-734.
- Mannion, A.F., & Dolan, P. (1996). Relationship between myoelectric and mechanical manifestations of fatigue in the quadriceps femoris muscle group. *European Journal of Applied Physiology and Occupational Physiology*, 74(5), 411-419.
- Mathur, S., Eng, J.J., & MacIntyre, D.L. (2005). Reliability of surface EMG during sustained contractions of the quadriceps. *Journal of Electromyography and Kinesiology*, 15(1), 102-110.
- Merletti, R., Lo Conte, L.R., & Orizio, C. (1991). Indices of muscle fatigue. *Journal of Electromyography and Kinesiology*, 1(1), 20-33.
- Portney, L.G., & Watkins, M.P. (2009). *Foundations of clinical research: Applications to practice*. Saddle River: Prentice-Hall.
- Rainoldi, A., Bullock-Saxton, J.E., Cavarretta, F., & Hogan, N. (2001). Repeatability of maximal voluntary force and of surface EMG variables during voluntary isometric contraction of quadriceps muscles in healthy subjects. *Journal of Electromyography and Kinesiology*, 11(6), 425-438.
- Rainoldi, A., Falla, D., Mellor, R., Bennell, K., & Hodges, P. (2008). Myoelectric manifestations of fatigue in vastus lateralis, medialis obliquus and medialis longus muscles. *Journal of Electromyography and Kinesiology*, 18(6), 1032-1037.
- Rogers, D.R., & MacIsaac, D.T. (2013). A comparison of EMG-based muscle fatigue assessments during dynamic contractions. *Journal of Electromyography and Kinesiology*, 23(5), 1004-1011.
- Santos, M.C.A., Semeghini, T.A., Azevedo, F.M., Colugnati, D.B., Negrão Filho, R.F., Alves, N., et al. (2008). Analysis of localized muscular fatigue in athletes and sedentary subjects through frequency parameters of electromyographic signal. *Brazilian Journal of Sports Medicine*, 14(6), 509-512.
- Silva, C.R., Geres, B.S., Kuriki, H.U., Negrão-Filho, R.F., Alves, N., & de Azevedo, F.M. (2012). Analysis of reliability of EMG signal frequency domain parameters used in the characterization of localized muscle fatigue. *Motriz: Revista de Educação Física*, 18(3), 456-464.
- Silva, C.R., Silva, D.O., Aragão, F.A., Ferrari, D., Alves, N., & Azevedo, F.M. (2014a). Influence of neuromuscular fatigue on co-contraction between vastus medialis and vastus lateralis during isometric contractions. *Kinesiology*, 46, 179-185.
- Silva, C.R., Silva, D.O., Ferrari, D., Negrão Filho, R.F., Alves, N., & Azevedo, F.M. (2014b). Exploratory study of electromyographic behavior of the vastus medialis and vastus lateralis at neuromuscular fatigue onset. *Motriz: Revista de Educação Física*, 20, 213-220.
- Watanabe, K., & Akima, H. (2010). Neuromuscular activation of vastus intermedius muscle during fatiguing exercise. *Journal of Electromyography and Kinesiology*, 20(4), 661-666.
- Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231-240.

Submitted: June 26, 2015

Accepted: January 31, 2016

Correspondence to:

Fábio Mícolis de Azevedo

Department of Physiotherapy FCT/UNESP

Roberto Simonsen Street, 305

19060-900, Presidente Prudente (SP), Brazil

Phone: 551896139152 / 551832295820

E-mail: micolis@fct.unesp.br

Acknowledgments

This research was supported by São Paulo Research Foundation (FAPESP), No 2008/11639-4.