A study of the reproducibility and reliability of the Musculo-Articular Stiffness of the ankle joint

Federico Paris-Garcia[a*], Alberto Barroso[b], Federico Paris[b]

[a]Department of Sports and Computers, University Pablo de Olavide, SEVILLE (SPAIN)

[b]School of Engineering, University of Seville, SEVILLE (SPAIN)

*Corresponding Author: Federico Paris-Garcia

Email: fparis@upo.es

Orcid Id: 0000-0002-7440-5719

Abstract

The objective of this work was to determine the reproducibility and reliability of the musculo-articular stiffness (MAS) of the ankle joint by the free vibration technique, and to evaluate its usefulness as well. Seventeen (nine males and eight females) healthy university students were included in the study. Force (f), MAS (k) and unitary MAS (k_u) (defined as the ratio between the value of stiffness k obtained in the test *(absolute terms)* and the value of force (f)) were obtained. A test-retest protocol was designed and performed on the same day to determine the short-term reproducibility of f, k and k_u. Short-term reproducibility of k and k_u on 1 day in absolute terms (<7% Coefficient of Variation (CV)) and relative reproducibility (Intraclass Correlation Coefficient (ICC) and Pearson ≥0.97) for both feet were obtained. The reliability of k and k_u in absolute terms (<9% CV) and in relative terms (ICC and Pearson ≥0.93) based on repeating the protocol for 1 week was analysed for both feet. To analyse the usefulness, the Effect Size (ES) ratio = "Trivial" for all variables (for 1 day and 1 week) and the Smallest Worthwhile Change (SWC) ratio (Typical Error (TE)<SWC) = "GOOD" for k and k_u (1 day and 1 week) were considered. The Minimal Difference (MD) needed to be considered "real" for k_u ≅3.5% (1 day); k_u≅8.5% (1 week) ($p<0.05$) was obtained. The statistical analysis carried out displayed the high reproducibility, reliability and usefulness of the MAS test, which was more consistent with k_u than k. Therefore, the unitary stiffness (k_u) proven to be representative of the mechanical response of the ankle joint obtained by free vibration techniques, which allows comparison between different subjects.

Keywords: Stiffness, functional response, free vibration technique, reliability.

1. INTRODUCTION

The use of stiffness to evaluate the mechanical behaviour of muscle-tendon units (MTUs) has been widely accepted in the scientific literature in past decades, as reported in the comprehensive review by Ditroilo et al. (2011b).

Generally, the concept of stiffness associated with the MTU establishes that, for a given applied load, smaller elongations imply stiffer MTU or less compliance.

A stiffer MTU has multiples advantages: for instance, it may be able to transmit contractile force to the skeletal segment more efficiently and rapidly (Walshe et al., 1996; Watsford et al., 2010). As a stiffer MTU exhibits more resistance to deformation, more energy must be applied, increasing its capacity to absorb vibrations, which are considered modulators of potential injury (Wilson et al., 1991). Moreover, the energy required in the stretch-shortening cycles (SSCs) of the involved muscles to develop a certain movement decreases as the stiffness increases (lower energy costs) (Lacour and Bourdin, 2015; París-García et al., 2013; Paris-Garcia, 2010). However, compliance of the MTU has also been identified to have some advantages during locomotion requiring different MTU properties (Biewener 1998, Mörl et al. 2016).

Therefore, the stiffness parameter is a significant factor related to muscle function (Wilson et al., 1994) and athletic performance (Heise and Martin, 1998; Walshe and Wilson, 1997), as has been shown in different proposals. The stiffness of the MTU is related to performance during fast and slow SSC movements (Chelly and Denis, 2001; Ditroilo et al., 2011a, 2011b; Walshe and Wilson, 1997; Wilson et al., 1994).

Various methods have been reported in the scientific literature to obtain the stiffness linked to the MTU in different parts of the body. The aim of the present work focused on the stiffness obtained by the application of free vibration techniques. Numerous methods to assess stiffness using single and multi-joint protocols based on various mathematical models have been used for the ankle joint (Ditroilo et al., 2011b; Faria et al., 2009; Fukashiro et al., 2001; Hunter and Spriggs, 2000; Kongsgaard et al., 2011; McLachlan et al., 2006; Murphy et al., 2003; París-García et al., 2013; Shorten, 1987). Other alternative methods to determine MTU properties can be found in (Penasso and Thaller 2017, Siebert et al. 2007). Also, it should be mentioned a pioneering study by Christensen et al. (2017) to measure the MTU stiffness without interfering influence of e.g. antagonistic muscles, skin, and articular capsule.

The applications of the above methods (free vibration techniques) to obtain the stiffness of the MTU gave rise to the consistent use of the term musculo-articular stiffness (MAS) (Ditroilo et al., 2011b) (henceforth *k*). MAS is a global measure of stiffness that incorporates not only the muscle-tendon structure but also skin, ligaments and articular surfaces (Rabita et al., 2008). Various assessments have demonstrated that MAS is a relevant parameter, as higher MAS values are associated with superior muscular performance (e.g., Murphy et al., 2003; Watsford et al., 2010; Wilson et al., 1994) and higher levels of functional capacity (Faria et al., 2009, 2010).

The *k* value can be obtained from a wide variety of methodologies, making the comparison of data among studies very difficult. Studies that have focused on obtaining the *k* of the ankle joint have utilized several experimental approaches. Some of them obtained the mechanical response of the ankle joint (traditionally linked to the triceps surae) using the free vibration technique (Babic and Lenarcic, 2004; Blackburn et al., 2006; Faria et al., 2009; Fukashiro et al., 2001; McLachlan et al., 2006; Murphy et al., 2003; París-García et al., 2013). Among them, only Babic and Lennarcic (2004) used a procedure based exclusively on the rotation of the foot around the ankle articulation. The remaining procedures were based on the vertical displacement of the lower leg (Blackburn et al., 2006; Faria et al., 2009; Fukashiro et al., 2001; McLachlan et al., 2006; Murphy et al., 2003; París-García et al., 2013).

To establish the representativeness of *k* (related to the ankle joint), different proposals have been developed related to the construct validity of free-oscillation techniques. Some proposals have related *k* to the rate of torque development (RTD), the ratio of maximum force developed (RFD) and electromechanical delay (EMD), (Ditroilo et al., 2011a; Watsford et al., 2010; Wilson et al., 1992). Among available research, strong reliability of a test of lower body stiffness was previously reported (Walshe et al., 1996). Furthermore, other proposals reported very good reliability for unilateral ankle stiffness (Murphy et al., 2003) and acceptable reliability for bilateral ankle stiffness (McLachlan et al., 2006).

All procedures considered above yield a k value from a force registered at a measuring device in one test. This implies that a subject with larger anthropometric characteristics or greater weight when placed on the measurement device will apply higher force at the load cell and will obtain higher values of k regardless of the level of fitness.

A new parameter, the unitary stiffness k_u, which is derived from the ratio between the value of stiffness k obtained in the test *(absolute terms)* and the value of force (f) registered at the measurement device for one subject (obtained in the same test), is defined in the present work. This parameter allows us to compare one subject in two different moments or two subjects with different heights or weights.

Because k obtained by these procedures in absolute terms is largely affected by the influence of other variables such as anthropometric measures, it is also necessary to analyse the behaviour of these other variables in test-retest protocols of this new parameter. This would allow us to make comparisons at different times and between different subjects.

Furthermore, the purpose of the current research is to use k_u as a parameter to discriminate between different subjects, and therefore, it is very interesting to evaluate the usefulness of the test. For this purpose, the statistical analyses carried out with other physical tests to obtain the performance of athletes (Buchheit et al., 2010; Ferrete et al., 2014) are considered. To evaluate the usefulness of this procedure, a comprehensive statistical analysis is developed in the methods section.

In summary, the aim of the present work was to evaluate the reproducibility, short-term reliability and usefulness of the entire process that would permit accurate unilateral assessment of k and k_u related to the ankle joint response based on the free vibration technique.

2. METHODS

2.1. Subjects

Seventeen healthy active university students (9 males and 8 females) [age (mean 23.13) (SD 2.85) years, mass (mean 68.69) (SD 14.20) kg, height (mean 174.81) (SD 9.57) cm] volunteered to participate in the current study. All subjects were medically screened to determine their health and exercise habits prior to testing and to ensure they did not have any previous injury to the lower body musculature. Prior to testing, all subjects attended a familiarization session which involved performing all test items, with particular attention to the lower body stiffness test. Each subject gave written informed consent to participate in the study, which was approved by the University Ethics Committee of University Pablo de Olavide and University of Seville.

2.2. Research design

To evaluate the reproducibility and reliability of the procedure, the subjects had to visit the laboratory twice with one week between visits. An identical protocol was followed to standardize any other effects. The sample size used in the reliability study was consistent with the sample sizes used in previous reliability studies related to obtaining musculo-articular stiffness (MAS) around the ankle articulation (Ditroilo et al., 2011b). Subjects were instructed to refrain from vigorous lower body exercise 48 h prior to each test day and required to maintain a constant routine. According to the study of Zinder et al. (2007), half of the body weight was considered to select the proper load in the test designed to obtain the MAS (k) at a fixed load. To obtain k and k_u, 2 consecutive tests on the same leg were carried out. The subject was familiarized with the protocol during the first visit to the laboratory. The data from the first and second days of the protocol and their differences were used in the present study of the reliability of the procedure. An identical procedure was carried out with the contralateral leg. The individual test duration was less than 5 seconds and sufficient recovery time between tests was given.

2.3. Test protocol

2.3.1. Warm-up

Participants warmed up by cycling at 100 W for 5 min maintaining cadence between 60-70 rpm. During this time, an explanation of the entire testing protocol was given.

2.3.2. Musculo-articular stiffness

The test considered in this work to obtain *k*, defined as MAS of the muscles linked to ankle articulation, was based on the free vibration technique (París-García et al., 2013). The response of the subject corresponded to that of a damped single degree of freedom (DOF) system and is associated with the vertical displacement of the shank linked to rotation of the ankle articulation.

The subject adopts a position in the test, see Fig. 1a, so that the lower body is capable of attenuating the vibration originated by a disturbance, thus assumed to act as a damped single DOF system. The disturbance is generated by the free fall of a mass, being the height and weight always the same, thus the impact energy is constant for all tests. This is importat as stated out by Kearney et al. 1982, Blackburn et al. 2008. This shock absorbing capacity is determined by the mechanical response of the muscles responsible for plantarflexion of the foot (the static situation is quasi-linearly related to the isometric force (Maton et al., 1987; Queisser et al., 1994)). In this proposal, the measuring device used was that developed in (París-García et al., 2013), where the subject places the forefoot area on the load cell, the ankle and the knee adopt a neutral position, and the hip is flexed at 90° (see Fig. 1).

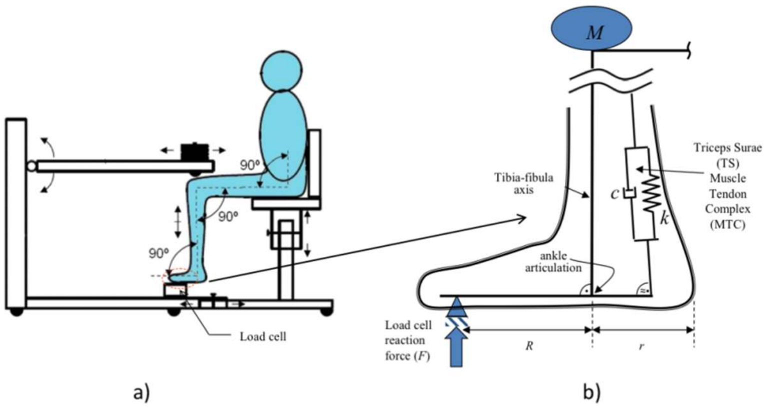


Fig. 1. a) Position of the subject for making measurements (mechanical properties of the ankle joint) (k) using the free vibration technique; b) schematic details of the balancing around the ankle joint.

2.4. Statistical analysis

Various statistical tools were used to evaluate the accuracy, the reliability and the usefulness of the present test.

The distribution of each variable was examined with the Shapiro-Wilk normality test. Homogeneity of variance was verified by the Levene test, and data herein are presented as the means and standard deviations (+ or - SD).

To examine the short-term reliability of the test over two consecutive trials, pairwise comparisons were first applied to determine any learning effect or systematic bias with Bland-Altman graphic analysis. The magnitude of differences between consecutive trials was expressed as the standardized mean difference (considering Cohen effect sizes, ES). The criteria to interpret the magnitude of the ES were as follows: <0.2 trivial, 0.2 to 0.5 small, 0.5 to 0.8 moderate, >0.8 large (Smith and Hopkins, 2011). The change in the mean value between trials (difference), the typical error of measurements (TE), the typical error expressed as a coefficient of variation (C.V., %), the intraclass correlation coefficient (ICC) and the Pearson coefficient were determined by using the spreadsheet of Hopkins et al. (2009a). The interday reliability was calculated from values obtained in different days. To establish the reliability of physical tests, earlier proposal, establish as reliable C.V. of mean value results lower than 5% (Cormack et al., 2008; Impellizzeri et al., 2008; Spencer et al., 2006).

The reliability of the main parameter k (MAS) has been established through several studies, but there is a need to establish the usefulness of k as a parameter to discriminate performance among subjects. In this sense, it is very interesting to analyse the sensitivity of other variables (e.g., fatigue, training…) to k. The usefulness of the test was assessed comparing the smallest worthwhile change (SWC; 0.2 multiplied by the between-subject deviation (SD), based on Cohen's effect size principle) with the typical error (Smith and Hopkins, 2011). If the typical error was below the SWC, the test was rated as "GOOD"; if the typical error was similar to the SWC, the test was rated as "OK"; and if the typical error was higher than the SWC, the test was rated "Marginal" (Hopkins, 2004). "Smallest difference needed to be considered real"

(MD; corresponding to a change likely to be “almost certain”) was calculated as TE × 1.96 × √2.17 (Weir, 2005).

For all analyses, the level of significance was set to $p<0.05$. For all pairwise comparisons, thresholds for clinical differences were calculated as 0.2 × the pooled SD of the two groups of interest (Myles and Cui, 2007).

3. RESULTS

3.1. Results associated with MAS (k)

Values for *f*, *k* across the test-retest protocol at one day and one week of assessment for both feet are presented in Fig. 2.

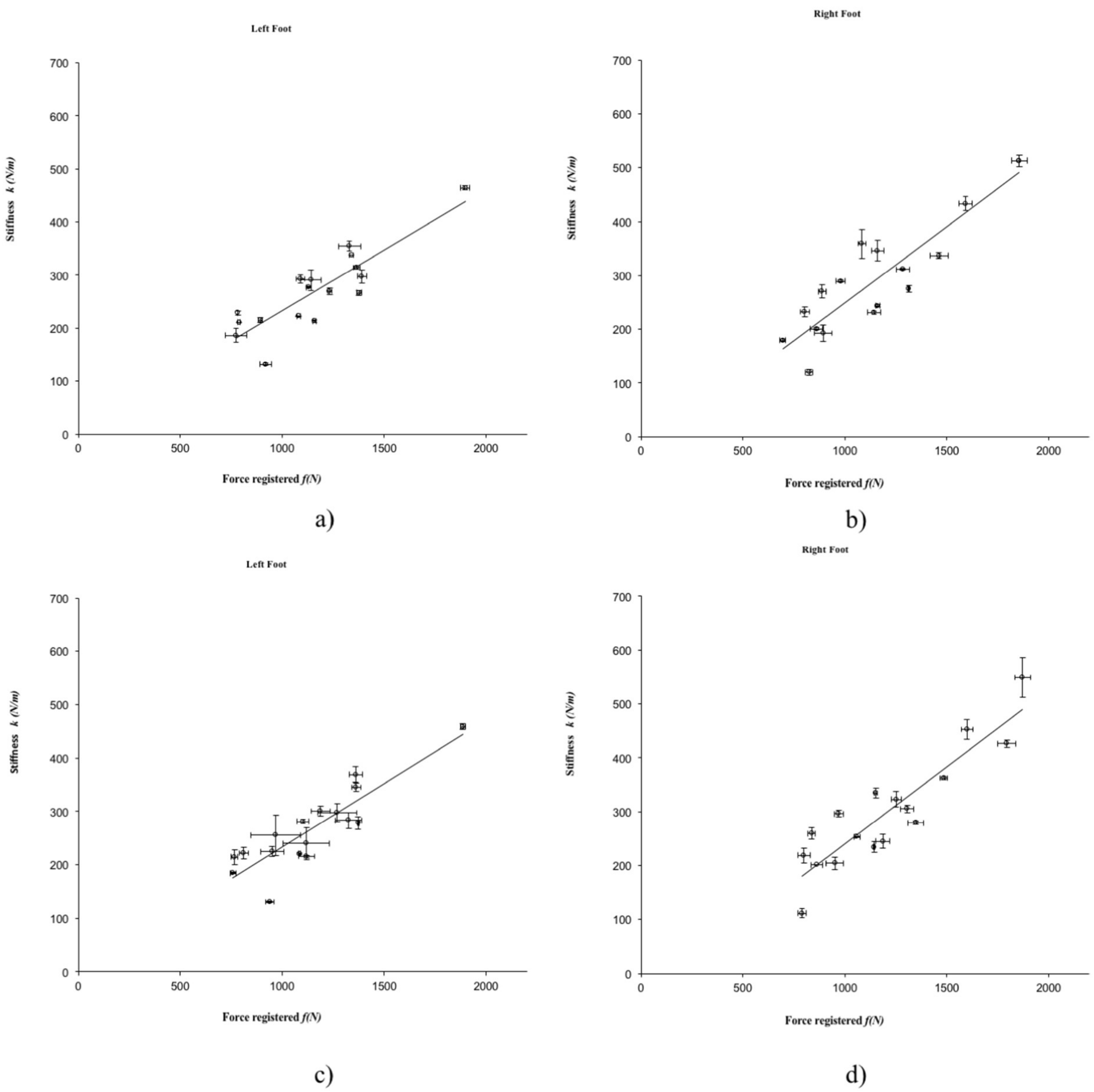


Fig. 2. Results of f and k obtained for test-retest applied in one day: a) Left foot, b) Right foot. The results of f and k obtained for test-retest applied in one week. c) Left Foot. d) Right Foot.

The x-axis displays the force *f* registered in the load cell, and the y-axis displays the obtained *k*. Each point represents a subject, and each side of the cross shows the difference in the values between test and retest for both variables.

The differences in *f* values registered between both tests (test-retest) for all subjects are small (smaller in the left foot than in the right foot, Fig. 2a. and 2b). Additionally, this variability in *f* values was associated with small variability in the values of *k* obtained for both feet.

This way of representing *k* from the registered *f* values allows the comparison between different subjects.

The location of each point of both variables (*f* and *k*) for each subject gives a visual idea of k_u. The position of each point to the experimental cloud and the trend line provide a visual representation of the stiffness of each subject in relation to the rest of the subjects.

In this sense, a subject located above the trend line is stiffer than a subject located under the trend line, considering the force registered for both subjects. In the case that both subjects are above the trend line, the subject farther from the trend line is stiffer than the other subject.

The results obtained from test-retest during one week are represented in Fig. 2c and 2.d. It can be observed that in several subjects, the results during one week with the left foot were slightly more variable than the test-retest results of one day, showing greater length of the "cross arm". This fact has no effect on the k_u variability because *f* and *k* change in the same way (as detailed in Fig. 3). A similar plot for parameter *k* would have a much higher dispersion for the parameter, which is, for all the above comments, not realistic.

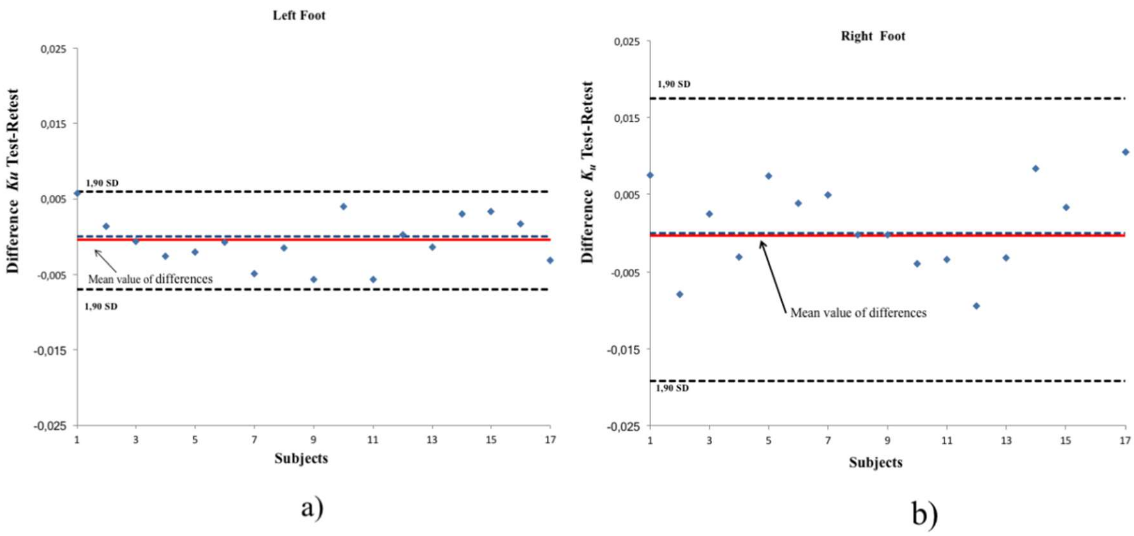


Fig. 3. Bland-Altmand graph of test-retest for the variable k_u: a) Left foot, b) Right foot.

Fig. 3 shows the level of short reproducibility of k_u for both feet and for all subjects. The Bland-Altman graphs analyse the level of agreement of two consecutives measures (e.g., test-retest) (Bland and Altman, 1986).

If the results of both the first and second test (test-retest) are the same, then the difference between them is zero. In the case that the first measure is higher than the second, the score will be positive, and vice versa.

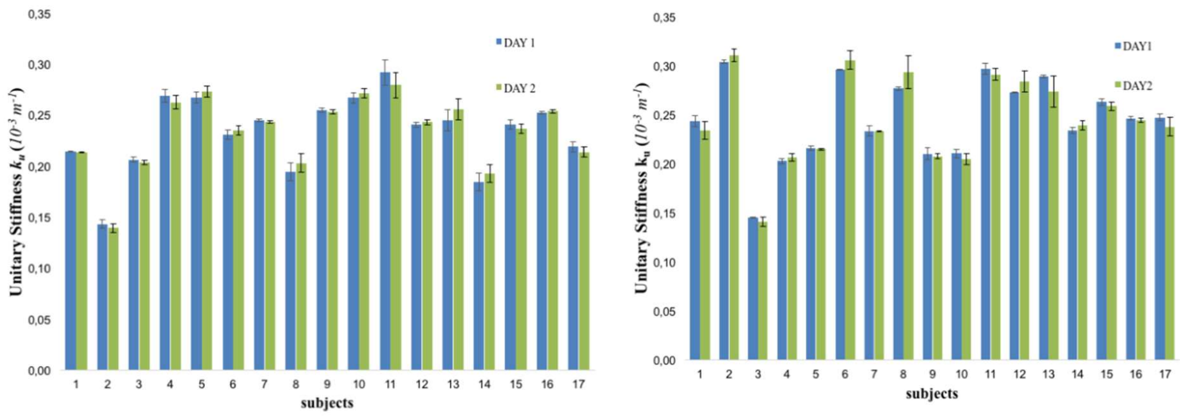
The mean value of the differences between both measures for all subjects and the 90% confidence limits as mean differences (1.9 SD) are plotted in Fig. 3. It is expected that the 95% limits include 90% of the differences between the two measurements.

In the left foot (LF), there is an unclear trend across the subjects of the sample, and the mean value is slightly lower than zero (see Fig. 3.a). All the differences along the set are inside the limits of agreement, and the limits are very narrow.

In the right foot (RF) (see Fig. 3.b), the trend in the differences is unclear along the set of subjects as well. This aspect shows that there is no learning effect between both measurements.

The mean value of the differences is close to zero, and the limits of agreement are wider than in the LF.

Fig. 4 shows the evolution in k_u values for all subjects during one week for the LF and the RF. The differences are not significant between feet even during one day or one week (p>0.05). The variability observed in the values of f and k is due to the variability of the subject placing his/her foot on the measuring device and has no effect in k_u values.



a) b)

Fig. 4. Comparison of values of k_u for one week for all subjects (left foot). The differences in the mean values are not significant for any subject (p>0.05). a) Left foot. b) Right foot.

3.2. *Results associated with reliability and usefulness of the test*

All the variables used to establish short-term reliability, interday reliability and usefulness are presented in Tables 1 and 2. The pairwise analysis revealed no significant differences between the two trials (all p>.05).

Table 1. Reliability measurements for the test-retest (1 day) protocol for force f, stiffness k, and unitary stiffness k_u.

Statistical variables	Force (f) (N)		Stiffness (k) (N/m)		Unitary Stiffness ku (k/f) (10-3·m-1)	
	Left Foot	**Right Foot**	**Left Foot**	**Right Foot**	**Left Foot**	**Right Foot**
Mean value	**1155.50**	**1189.19**	**267.45**	**291.24**	**0.232**	**0.246**
Standard deviation	286.71	307.86	74.73	91.12	0.035	0.042
(number of subjects)	17	17	17	17	17	17
TE (90% CL)	**37.36**	**38.13**	**11.86**	**17.58**	**0.007**	**0.008**
(upper limit)	52.96	54.06	16.81	24.92	0.010	0.011
(lower limit)	29.1	29.7	9.3	13.7	0.005	0.006
Typical error as C.V. (90% CL)	**3.9**	**3.6**	**5.1**	**6.0**	**3.0**	**3.3**

(upper limit)	5.7	5.2	7.5	8.7	4.3	4.7
(lower limit)	3.0	2.7	3.9	4.7	2.3	2.6
Difference (90% CL)	**4.93**	**14.65**	**2.03**	**9.41**	**0.001**	**0.001**
(upper limit)	27.30	37.49	9.14	19.94	0.132	0.006
(lower limit)	-17.44	-8.19	-5.07	-1.12	-0.003	-0.003
ES (90% CL)	**0.01**	**0.03**	**0.01**	**0.04**	**0.014**	**0.005**
(upper limit)	0.09	0.10	0.10	0.13	0.165	0.091
(lower limit)	-0.07	-0.05	-0.08	-0.06	-0.138	-0.081
(ratio)	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL
(p>0.05)	0.86	0.56	0.85	0.49	0.876	0.919
SWC	**57.3**	**61.6**	**14.95**	**18.2**	**0.007**	**0.008**
(%)	4.96	5.18	5.59	6.26	3.00	3.41
(ratio)	**GOOD**	**GOOD**	**GOOD**	**GOOD**	**GOOD**	**GOOD**
MD (90% CL)	**73.22**	**74.74**	**23.25**	**34.46**	**0.013**	**0.016**
(%)	4.96	5.18	5.59	6.26	3.00	3.41
ICC (90%)	**0.99**	**0.99**	**0.98**	**0.97**	**0.97**	**0.97**
(upper limit)	0.99	0.99	0.99	0.99	0.99	0.99
(lower limit)	0.97	0.97	0.95	0.94	0.93	0.93
Pearson (90% CL)	**0.98**	**0.99**	**0.98**	**0.98**	**0.97**	**0.97**
(upper limit)	0.99	1.00	0.99	0.99	0.99	0.99
(lower limit)	0.96	0.98	0.95	0.96	0.93	0.92

Note: Typical error of measurement (TE), TE expressed as a coefficient of variation (C.V.), difference in the mean between the two trials, effect size (ES) and ES rating (see Methods), smallest worthwhile change (SWC) and rating of the test (see Methods), minimal difference needed to be considered "real" (MD)

Table 2. Reliability measurements for the test-retest (1 week) protocol for force *f*, stiffness *k*, and unitary stiffness k_u

Statistical variables	Force (f) (N)		Stiffness (k) (N/m)		Unitary stiffness ku (k/f) (10-3·m-1)	
	Left Foot	**Right Foot**	**Left Foot**	**Right Foot**	**Left Foot**	**Right Foot**
Mean value	**1118.70**	**1167.29**	**261.00**	**286.58**	**0.237**	**0.243**
Standard deviation	281.32	282.88	77.10	100.17	0.037	0.440
(number of subjects)	17	17	17	17	17	17
TE (90% CL)	**74.14**	**48.23**	**20.54**	**18.11**	**0.0072**	**0.0104**
(upper limit)	105.10	68.37	29.12	25.07	0.010	0.015
(lower limit)	57.8	37.6	16.0	14.3	0.006	0.008
Typical error as C.V. (90% CL)	**7.2**	**4.7**	**8.5**	**6.9**	**3.1**	**3.9**
(upper limit)	10.3	3.7	12.3	9.7	4.5	5.6
(lower limit)	5.6	6.8	6.6	5.4	2.4	3.0
Difference (90% CL)	**39.27**	**2.65**	**7.47**	**5.89**	**-0.002**	**0.003**
(upper limit)	83.66	31.53	19.77	16.08	0.002	0.010
(lower limit)	-5.13	-26.23	-4.83	-4.30	-0.007	-0.003
ES (90% CL)	**0.01**	**0.00**	**0.09**	**0.02**	**-0.054**	**0.045**
(upper limit)	0.09	-0.08	0.25	0.14	0.044	0.236
(lower limit)	-0.07	0.08	-0.06	-0.10	-0.152	-0.147
(ratio)	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL	TRIVIAL
(p>0.05)	0.86	0.94	0.30	0.75	0.351	0.693
SWC	**56.3**	**56.6**	**15.42**	**20.0**	**0.0073**	**0.0880**
(%)	5.03	4.85	5.91	6.99	3.09	4.85

(ratio)	**MARGINAL**	**GOOD**	**GOOD**	**GOOD**	**GOOD**	**GOOD**
MD (90% CL)	**145.32**	**94.52**	**40.25**	**35.49**	**0.014**	**0.020**
(%)	12.99	8.10	15.42	12.38	5.99	8.41
ICC (90%)	**0.94**	**0.98**	**0.94**	**0.97**	**0.97**	**1.00**
(upper limit)	0.97	0.99	0.97	0.99	0.99	1.00
(lower limit)	0.87	0.95	0.86	0.94	0.92	1.00
Pearson (90% CL)	**0.93**	**0.98**	**0.93**	**0.97**	**0.96**	**0.96**
(upper limit)	0.97	0.99	0.97	0.99	0.98	0.98
(lower limit)	0.84	0.95	0.84	0.94	0.92	0.91

Note: Typical error of measurement (TE), TE expressed as a coefficient of variation (C.V.), difference in the mean between the two trials, effect size (ES) and ES rating (see Methods), smallest worthwhile change (SWC) and rating of the test (see Methods), minimal difference needed to be considered "real" (MD)

The mean values and the statistical analysis of *f*, *k* and k_u for both legs are represented in the three main columns of Tables 1 and 2. In the first row, the mean value, standard deviation (SD) and numbers of subjects are represented.

In the second and third rows, low typical errors are presented, and typical errors expressed as coefficient of variation (C.*V*.) are lower than 6% of the mean values for all variables and for both feet.

The differences in all indices between repeated tests show trivial differences related to the effect size *(ES),* which is close to zero with confidence limits between -0.2 to 0.2 (Hopkins et al., 2009a, 2009b). Values of the smallest worthwhile change (SWC) are slightly higher for *f* and *k* than for k_u, but in any case, all are lower than 6.5% of the mean value. The *TE* is lower than *SWC* related to *f* and *k* (for both feet).

The values of TE are lower than the values of SWC for *f*, *k* and k_u (for both feet), and therefore, the ratio of the test is "GOOD" (Hopkins et al., 2009b) for all of them. The intraclass coefficient correlation (ICC) and Pearson coefficient shows values very similar with a very narrow confidence limit (0.93-0.9). All were higher than 0.96.

The results obtained from repeated tests along one week present slightly higher variability than the test-retest in one day (see Table 2). This fact only affects *f* and *k* values. Typical errors such as C.V. of *f* represent the differences in the position of the subject in the test along one week., the values for the LF are approximately 7%, and the values for the RF are lower than 5% of the mean value. Typical errors as C.V. values associated with *k* are linked to *f* values and are between 5 to 10% of the mean value for both feet.

From all results, the variability in *k* values is the least suitable for comparison even though they are below 8.6% between different days.

The best values of *SWC* are associated with k_u (3.09% LF; 4.18% RF).

The values of all variables present trivial differences associated with *ES*, showing values close to zero. The comparison between *TE* and *SWC* along one week provides a satisfactory ratio for the RF (*f*, *k*, and k_u). For the LF, the ratio is "marginal" to *f,* but the ratio is "GOOD" to *k* and k_u.

The values of ICC and Pearson along one week were very similar between both feet and close to the unit (>0.92). Over time, the reliability of all variables did not markedly change. In the same way, the best values of ICC and Pearson are linked to k_u for both feet (≥0.96).

4. DISCUSSION

The investigation carried out in the present work evaluates the short reproducibility, the interday reliability and the usefulness of proposed test to obtain the musculo-articular stiffness (MAS) and the unitary MAS of the ankle linked to the ankle joint response using the free vibration technique. The results of variables such as *f* and *k* (and their evolution) according to a test-retest protocol for one day and one week were analysed. Additionally, a different parameter, unitary stiffness k_u (which has a relevance that is not always highlighted in previous works) and its evolution according to the same protocol are also proposed and evaluated.

As a first observation, the information shown in Fig. 2 allows us to easily interpret and compare the values obtained from a set of subjects. Each test yields a pair of values, *f* and *k*. Several things can be visually observed: i) The obtained *k* is not a universal characteristic of humans. Different subjects with different characteristics have different values of *k* in addition to different values of *f*. ii) There is a small variability in the force registered *f* across to the different tests, which is linked to the small variability in *k*. A small variability in both variables is obtained for one day and one week. iii) Higher values of *k* do not imply for itself, interpretation associated with the level of performance. iv) By adding the linear trend line of the experimental cloud, it is possible to easily compare among different subjects by measuring the distance between each point to the mentioned line. For instance, a taller subject with a greater weight will register a higher value of *f* and will yield higher *k* values, but the stiffness compared with another subject is represented by the differences in the distances from the various points of both subjects to the trend line.

Fig. 2 shows that the slope of the trend line is higher in the RF than in the LF for the set of subjects. Therefore, from a given force value, higher stiffness values are obtained in the LF than in the RF.

On one hand, regarding the reproducibility (absolute reliability as C.V. and relative reliability as ICC defined by Ditroilo (2011b)) of *f* and *k*, it can be observed LF has larger differences between the *f* values of test and retest than RF (3.9% C.V. (LF) and 3.6% C.V.(RF)). On the other hand, *k* values show slightly higher differences (test-retest) in the RF (6% C.V.) than in the LF (5.1% C.V.) (see Table 1). This fact can be observed in Fig. 2, which shows differences in the length of the crosses for all subjects. The variability obtained in *f* values is due to the variability in the position of the whole leg in the test and does not imply higher variability in the *k* values (≅5% of C.V.). The results of the present study are in accordance with the results obtained in the literature (<10% of C.V.) (Murphy et al., 2003; Walshe et al., 1996).

The relative reliability according to the ICC and Pearson coefficient show values and confidence limits very similar to those of *f* and *k* (>0.97). They are considerably higher than those shown in the first reported studies (Hunter and Spriggs, 2000; McNair and Stanley, 1996; Walshe et al., 1996) but similar to those reported in the most recent studies (McLachlan et al., 2006; Murphy et al., 2003). The k_u variable showed very high agreement between test and retest for both feet (Fig. 3). There is no clear trend allowing it to be established whether the retest value overestimated or underestimated the test value (e.g., learning effect). In any case, all the difference values are between the confidence limits of agreement for all subjects, and the mean value of the difference is close to zero for both feet. The k_u variable shows extremely high values of relative reliability and absolute reliability for both feet (C.V.<3.5% and ICC=0.97).

One explanation may be that f and k change in the same sense. Therefore, this parameter is highly suitable as a reference variable to perform comparisons.

According to Fig. 4 and the values of f and k in Table 2, the interday reliability with time is slightly higher in the LF than in the RF (C.V. of (f_{LF}=7.2%, f_{RF}=4.7%; k_{LF}=8.5%, k_{RF}=6.9%)). In the LF for several subjects, the very small variability in f implies a slightly higher difference in k obtained with time. In the RF, for all subjects, the small variability in f also implies a small variability in k. The differences in f between both tests with time are clearly due to the position of the subject.

k_u follows the same trend with time as the test-retest in one day. While the C.V. values of f and k are higher for the LF than for the RF, the C.V. values of k_u are lower for the LF than for the RF. The absolute reliability of k_u obtained in the present study (C.V.<4%) is in accordance with the best scores obtained in previous studies related to mechanical responses (Ditroilo et al., 2011b; Watsford et al., 2010) close to 4% of C.V. The evolution of the mean values of k_u during one week show no significant difference between day one and day two (p>0.05) for both feet (see Fig. 4).

To be used as an experimental test of functional responses such as usefulness, the ratio determined by the test must be classified as “GOOD” or, at minimum, “OK” (Buchheit et al., 2010). In this sense, the TE must be lower than the SWC. It means that the noise of the signal (response measured from the subject) is lower than the minimal detectable change according to the size of the set. The obtained TE and SWC values allow the test to be rated as “GOOD” for all variables (f, k and k_u) and for measures made in the same day. The ratio of the test is “MARGINAL” for the LF associated with f. As mentioned before, these differences in f are due to aspects associated with the experimental procedure. Therefore, it is extremely important to monitor the entire process. Additionally, this influence of the variability in f is minimized when k_u is taken as a reference value. The usefulness of the test measurement over time is rated as “GOOD” for k and k_u for both feet (see Table 2).

Finally, from a practical point of view, the changes in the variables as meaningful or “almost certain” based on the minimal difference needed to be considered “real” (MD) as the result of the influence of other variables (e.g., effect of fatigue) are considered. It is therefore suggested that k_u changes by at least 3.41% to measure the effect of other variables in the same day (cross-sectional studies) and by 8.41% to measure the effect of others variables over time (longitudinal studies). The results of MD related to k are less than 5.6% in the same day for both feet (see Table 1) and ≅15% for one week. This implies that for a cross-sectional study, the influence of other variables will be detected when k or k_u changes by more than 7-9%, which is very narrow. To analyse the influence of other variables on the mechanical response of ankle joint with time, k is less useful because it must change by approximately 15%.

5. CONCLUSIONS

In summary, the analysis proposed for the test in the present study to obtain the muscular-articular stiffness associated with the ankle articulation has a high short-term reproducibility and is reliable with time for healthy active of university students. The parameter k_u is traditionally less used in the literature, presenting a certain benefit over k that allows comparisons between different subjects regardless of anthropometric factors, such as weight or height. Moreover, this parameter is less sensitive to the changes introduced by the variability of the forces registered between different tests (due to the changes in the position of the subject) with lower C.V. values in one day and in one week.

Finally, as a result of the present work, it is recommended that k_u must change by at least 3.5% between tests performed on the same day or by 8.5% between tests performed on different days to discriminate the effect of the application of other variables (e.g., fatigue, performance or sports injury rehabilitation).

Acknowledgment

This research was supported by the Group of Elasticity and Strength of Materials (University of Seville), with the cooperation of the Faculty of Health Sciences of University of Seville and the Andalusian Sport Medicine Centre of Seville. Additionally, this study was supported by the faculty of Sport of University Pablo of Olavide, which provided all necessary facilities for carrying out the research.

Conflict of interest

None of the authors of this manuscript entitled “A study of the reproducibility and reliability of the Musculo-Articular Stiffness of the ankle joint" have any conflicts of interest.

References

Babic, J., Lenarcic, J., 2004. In vivo determination of triceps surae muscle-tendon complex viscoelastic properties. Eur. J. Appl. Physiol. 92, 477–484.

Blackburn, J.T., Padua, D.A., Weinhold, P.S., Guskiewicz, K.M., 2006. Comparison of triceps surae structural stiffness and material modulus across sex. Clin Biomech (Bristol, Avon) 21, 159–167. doi:10.1016/j.clinbiomech.2005.08.012.

Bland, J.M., Altman, D.G., 2010. Statistical methods for assessing agreement between two methods of clinical measurement. Int. J. Nurs. Stud. 47, 931–936.

Buchheit, M., Spencer, M., Ahmaidi, S., 2010. Reliability, usefulness, and validity of a repeated sprint and jump ability test. Int. J. Sports Physiol. Perform. 5, 3–17.

Chelly, S.M., Denis, C., 2001. Leg power and hopping stiffness: relationship with sprint running performance. Med. Sci. Sports Exerc. 33, 326–333.

Cormack, S.J., Newton, R.U., McGuigan, M.R., Doyle, T.L.A., 2008. Reliability of measures obtained during single and repeated countermovement jumps. Int. J. Sports Physiol. Perform. 3, 131–144.

Ditroilo, M., Watsford, M., De Vito, G., 2011a. Validity and inter-day reliability of a free-oscillation test to measure knee extensor and knee flexor musculo-articular stiffness. J. Electromyogr. Kinesiol. 21, 492–498. doi:10.1016/j.jelekin.2010.11.004.

Ditroilo, M., Watsford, M., Murphy, A., De Vito, G., 2011b. Assessing musculo-articular stiffness using free oscillations: theory, measurement and analysis. Sports. Med. 41, 1019–1032. doi:10.2165/11591470-000000000-00000.

Faria, A., Gabriel, R., Abrantes, J., Brás, R., Moreira, H., 2009. Triceps-surae musculotendinous stiffness: relative differences between obese and non-obese postmenopausal women. Clin Biomech (Bristol, Avon) 24, 866–871. doi:10.1016/j.clinbiomech.2009.07.015.

Faria, A., Gabriel, R., Abrantes, J., Brás, R., Moreira, H., 2010. The relationship of body mass index, age and triceps-surae musculotendinous stiffness with the foot arch structure of postmenopausal women. Clinical Biomechanics 25, 588–593. doi:10.1016/j.clinbiomech.2010.02.014.

Ferrete, C., Requena, B., Suarez-Arrones, L., de Villarreal, E.S., 2014. Effect of strength and high-intensity training on jumping, sprinting, and intermittent endurance performance in prepubertal soccer players. J. Strength Cond. Res. 28, 413–422. doi:10.1519/JSC.0b013e31829b2222.

Fukashiro, S., Noda, M., Shibayama, A., 2001. In vivo determination of muscle viscoelasticity in the human leg. Acta Physiol. Scand. 172, 241–248. doi:10.1046/j.1365-201x.2001.00866.x.

Heise, G.D., Martin, P.E., 1998. "Leg spring" characteristics and the aerobic demand of running. Med. Sci. Sports Exerc. 30, 750–754. doi:10.1097/00005768-199805000-00017.

Hopkins, W.G., 2004. How to interpret changes in an athletic performance test. Sportscience 8, 1–7.

Hopkins, W.G., Marshall, S.W., Batterham, A.M., Hanin, J., 2009a. Progressive statistics for studies in sports medicine and exercise science. Med. Sci. Sports Exerc. 41, 3–13. doi:10.1249/MSS.0b013e31818cb278.

Hopkins, W.G., Marshall, S.W., Batterham, A.M., Hanin, J., 2009b. Progressive statistics for studies in sports medicine and exercise science. Med. Sci. Sports Exerc. 41, 3–13. doi:10.1249/MSS.0b013e31818cb278.

Hunter, D.G., Spriggs, J., 2000. Investigation into the relationship between the passive flexibility and active stiffness of the ankle plantar-flexor muscles. Clin. Biomech. 15, 600–606.

Impellizzeri, F.M., Rampinini, E., Castagna, C., Bishop, D., Ferrari Bravo, D., Tibaudi, A., Wisloff, U., 2008. Validity of a repeated-sprint test for football. Int. J. Sports Med. 29, 899–905.

Kongsgaard, M., Nielsen, C.H., Hegnsvad, S., Aagaard, P., Magnusson, S.P., 2011. Mechanical properties of the human Achilles tendon, in vivo. Clin Biomech (Bristol, Avon) 26, 772–777. doi:10.1016/j.clinbiomech.2011.02.011.

Lacour, J.R., Bourdin, M., 2015. Factors affecting the energy cost of level running at submaximal speed. Eur. J. Appl. Physiol. 115, 651–673. doi:10.1007/s00421-015-3115-y.

Maton, B., Pérès, G., Landjerit, B., 1987. Relationships between individual isometric muscle forces, EMG activity and joint torque in monkeys. Eur. J. Appl. Physiol. Occup. Physiol. 56, 487–494. doi:10.1007/BF00417781.

McLachlan, K.A., Murphy, A.J., Watsford, M.L., Rees, S., 2006. The interday reliability of leg and ankle musculotendinous stiffness measures. J. Appl. Biomech. 22, 296–304.

McNair, P.J., Stanley, S.N., 1996. Effect of passive stretching and jogging on the series elastic muscle stiffness and range of motion of the ankle joint. Br. J. Sports Med. 30, 313–7, discussion 318.

Murphy, A.J., Watsford, M.L., Coutts, A.J., Pine, M.J., 2003. Reliability of a test of musculotendinous stiffness for the triceps-surae. Phys. Ther. Sport 4, 175–181. doi:10.1016/S1466-853X(03)00077-4.

Myles, P.S., Cui, J., 2007. Using the Bland-Altman method to measure agreement with repeated measures. Br. J. Anaesth. 99, 309–311. doi:10.1093/bja/aem214.

Paris-Garcia, F., 2010. In-vivo determination of the viscoelastic properties of the triceps surae by means of the free vibration technique. University of Seville.

París-García, F., Barroso, A., Cañas, J., Ribas, J., París, F., 2013. A critical study on the experimental determination of stiffness and viscosity of the human triceps surae by free vibration methods. Proc. Inst. Mech. Eng. H 227, 935–954. doi:10.1177/0954411913487851.

Queisser, F., Blüthner, R., Bräuer, D., Seidel, H., 1994. The relationship between the electromyogram-amplitude and isometric extension torques of neck muscles at different positions of the cervical spine. Eur. J. Appl. Physiol. Occup. Physiol. 68, 92–101. doi:10.1007/BF00599248.

Rabita, G., Couturier, A., Lambertz, D., 2008. Influence of training background on the relationships between plantarflexor intrinsic stiffness and overall musculoskeletal stiffness during hopping. Eur. J. Appl. Physiol. 103, 163–171. doi:10.1007/s00421-008-0679-9.

Shorten, M.R., 1987. Muscle elasticity and human performance. Med. Sport Sci. 25, 1–18.

Smith, T.B., Hopkins, W.G., 2011. Variability and predictability of finals times of elite rowers. Med. Sci. Sports Exerc. 43, 2155–2160. doi:10.1249/MSS.0b013e31821d3f8e.

Spencer, M., Fitzsimons, M., Dawson, B., Bishop, D., Goodman, C., 2006. Reliability of a repeated-sprint test for field-hockey. J. Sci. Med. Sport 9, 181–184.

Walshe, A.D., Wilson, G.J., 1997. The influence of musculotendinous stiffness on drop jump performance. Can. J. Appl. Physiol. 22, 117–132.

Walshe, A.D., Wilson, G.J., Murphy, A.J., 1996. The validity and reliability of a test of lower body musculotendinous stiffness. Eur. J. Appl. Physiol. Occup. Physiol. 73, 332–339.

Watsford, M., Ditroilo, M., Fernández-Peña, E., D'Amen, G., Lucertini, F., 2010. Muscle stiffness and rate of torque development during sprint cycling. Med. Sci. Sports Exerc. 42, 1324–1332. doi:10.1249/MSS.0b013e3181ce509d.

Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Cond. Res. 19, 231–240. doi:10.1519/15184.1.

Wilson, G.J., Elliott, B.C., Wood, G.A., 1992. Stretch shorten cycle performance enhancement through flexibility training. Med. Sci. Sports Exerc. 24, 116–123.

Wilson, G.J., Murphy, A.J., Pryor, J.F., 1994. Musculotendinous stiffness: its relationship to eccentric, isometric, and concentric performance. J. Appl. Physiol. 76, 2714–2719.

Wilson, G.J., Wood, G.A., Elliott, B.C., 1991. The relationship between stiffness of the musculature and static flexibility: an alternative explanation for the occurrence of muscular injury. Int. J. Sports Med. 12, 403–407. doi:10.1055/s-2007-1024702.

Zinder, S.M., Granata, K.P., Padua, D.A., Gansneder, B.M., 2007. Validity and reliability of a new in vivo ankle stiffness measurement device. J. Biomech. 40, 463–467. doi:10.1016/j.jbiomech.2005.12.007.

Kearney et al. (1982) Dynamics of human ankle stiffness: variation with displacement amplitude. J Biomech; 15 (10): 753-6

Blackburn et al. (2008) Muscle stiffness and spinal stretch reflex sensitivity in the triceps surae. J Athl Train 2008; 43 (1): 29-36

Biewener (1998) Muscle function in vivo: a comparison of muscles used for elastic energy savings versus muscles used to generate mechanical power. Am Zool 38:703-717

Mörl et al. (2016). Contraction dynamics and function of the muscle-tendon complex depend on the muscle fibre-tendon length ratio: a simulation study. Biomech Model Mechanobiol, 15(1), 245-258.

Penasso and Thaller (2017). Determination of individual knee-extensor properties from leg extensions and parameter identification. Mathematical and Computer Modelling of Dynamical Systems, 23(4), 416-438.

Siebert et al. (2007). An improved method to determine neuromuscular properties using force laws - From single muscle to applications in human movements. Hum Mov Sci, 26(2), 320-341.

Christensen et al. (2017). Strain in shock-loaded skeletal muscle and the time scale of muscular wobbling mass dynamics. Sci Rep, 7(1), 13266.