9-30-2020

Effects of an 8-week resistance training intervention on plantar flexor muscle quality and functional capacity in older women: A randomised controlled trial

Pedro Lopez
Edith Cowan University

Brendan James Crosby
Edith Cowan University

Bruna Patrícia Robetti

Douglas Jean Preussler Turella

Thaís Andréia Schepa Weber

See next page for additional authors

This Journal Article is posted at Research Online.
https://ro.ecu.edu.au/ecuworkspost2013/8750

Authors

Pedro Lopez, Brendan James Crosby, Bruna Patrícia Robetti, Douglas Jean Preussler Turella, Thaís Andréia Schepa Weber, Morgana Lima de Oliveira, and Anderson Rech

Experimental Gerontology

EFFECTS OF AN 8-WEEK RESISTANCE TRAINING INTERVENTION ON PLANTAR FLEXOR MUSCLE QUALITY AND FUNCTIONAL CAPACITY IN OLDER WOMEN: A RANDOMISED CONTROLLED TRIAL
--Manuscript Draft--

Corresponding Author:	Pedro Lopez Edith Cowan University Perth, Western Australia AUSTRALIA
First Author:	Pedro Lopez
Order of Authors:	Pedro Lopez
	Brendan Crosby
	Bruna Robetti
	Douglas Turella
	Thaís Weber
	Morgana Oliveira
	Anderson Rech

Abstract:	The present study examined 8 weeks of resistance training and its effects on muscle quality measures, plantar flexor muscle strength, muscle thickness and functional capacity in older women. Moreover, we tested if changes in muscle quality were associated with functional capacity. Twenty-four older women (66.3 ± 5.8 yrs; 69.0 ± 3.0 kg; 25.3 ± 1.4 kg.m-2) were recruited to the study. After completion of the baseline assessment, participants were randomly assigned to either the resistance training (RET, n= 12) or an active control group (CTR, n= 12). Muscle quality was evaluated through muscle echo intensity (MQEI) and specific tension (MQST). Muscle thickness, unilateral plantar flexor muscle strength and functional tests were evaluated at baseline and after the training period. After 8 weeks, both MQEI and MQST did not respond to the intervention. Furthermore, significant changes in stair climb performance (P<0.05) were not associated with plantar flexor-derived muscle quality (P>0.05). Finally, significant gains in muscle hypertrophy were observed in the RET group (P<0.01), while muscle strength failed to change significantly (P>0.05). In conclusion, a resistance training program provided significant benefits in the stair climb test, unrelated to plantar flexor-derived muscle quality measures as previously demonstrated in quadriceps femoris.

Skeletal muscle echo intensity and the notion of muscle quality are novel research concepts that are being reported in the literature on a more frequent basis. Several previous studies have reported significant reductions in echo intensity in response to resistance training in both younger and older adults. However, many previous investigators have focused their attention on the quadriceps femoris muscles. As such, the purpose of the present study was to examine the effects of a eight week resistance training intervention on echo intensity, muscle quality (specific tension), and functional outcomes in older adults, with a secondary purpose of examining the associations among change scores. I believe that this study is of interest to the research community and is an excellent fit for Experimental Gerontology. However, I have several significant concerns that need to be addressed prior to acceptance. I have provided my comments below in chronological order separated into major and minor comments. Intervention studies such as the present one are very challenging to carry out, so I commend the authors for their efforts.

Answer: We really appreciate the positive and constructive comments from the reviewer. Thank you very much.

MAJOR COMMENTS:
1) One of my main concerns about the study was whether the training volume for the calf musculature was sufficient to elicit meaningful changes in muscle morphology. If the authors' emphasis was on the calf musculature, why were exercises for the upper limbs included with so few exercises that directly targeted the tested muscles? This needs to be considered and deliberated in the discussion section.

Answer: Thank you for the comment. This question has been very important during and after the conception of our study. We would like to share some thoughts on this.

First, previous studies have prescribed one (e.g., leg extension) (Radaelli et al., 2019, PMID: 29730331) or two resistance exercises (e.g., leg press and leg extension) (Radaelli et al., 2014, PMID: 24414336; Wilhelm et al., 2014, PMID: 25449853) for quadriceps femoris adaptations. These studies reported significant changes (~5 to 20%; P< 0.05) on muscle quality using single or multiple sets (Radaelli et al., 2019, PMID: 29730331; Radaelli et al., 2014, PMID: 24414336; Wilhelm et al., 2014, PMID: 25449853). Thus, if single sets of one or two resistance exercises were sufficient to stimulate changes in quadriceps femoris echo intensity, a prescription involving standing calf raises, 3 sets of 6-12 repetitions at 8-15RM was hypothesised to be equally capable of providing similar changes in calf muscles. However, differences were not observed for 8 weeks, suggesting that calf muscles may need greater local stimulus. Thus, our resistance training prescription design was in accordance with previous literature but produced different results than expected.

We have addressed this issue at:
Pages 3, lines 79-82: "Thus, different exercise modes, or even lower volumes of resistance training (i.e., single sets, or 1-2 resistance exercises) can promote changes to quadriceps derived MQEI by non-contractile tissue reduction as suggested by the authors (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b)."

Page 15, lines 345-349: "Thirdly, although quadriceps femoris may respond to a relatively low exercise stimulus as observed in previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014), the proposed number of exercises, intensity or volume (standing calf raise, 3 sets of 6-12 repetitions at 8-15RM) might have been an insufficient stimulus for plantar flexor muscles based on this muscle groups level of activity.".

Moreover, during the conception of this trial, we believed that a general resistance training program would be in accordance with exercise guidelines for older adults (Fragala et al., 2019, PMID: 31343601) and thereby, more appropriate than 2-3 calf specific resistance exercises. We were interested in providing all the possible benefits exercise affords older adults, and not just those tested by ourselves.

In order to make it clearer, we provided changes at:
Page 18, lines 440-447: "Finally, our resistance training prescription did not solely target the plantar flexor muscles and may not have provided a stimulus sufficient to improve MQEI or functional test results. However, previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b) demonstrated that quadriceps femoris MQEI was likely to present changes in response to single sets of one or two resistance exercises. With this in mind, we targeted multiple muscle groups in accordance with the latest exercise guideline for older adults, focusing on the overall benefits for functional capacity (Fragala et al., 2019) and not exclusively those tested by ourselves."

2) The authors' approach to the statistical analysis could use revision. I have three main qualms.
First, the authors' primary tool for examining both changes between and withing groups should be effect sizes, followed by 95% confidence intervals. The authors' reliance on p values and NHST is not in line with current recommendations. Please see:
https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.nature.com%2Farticles%2Fd41586-019-00857-
9&data=02%7C01%7Cplopezda%40our.ecu.edu.au%7C99484899c5ba44b5982108d7f579daee%7C9bcb323d7fa345e7a36f6d9cfdbcc272%7C1%7C0%7C637247776028988378&sdata=p6FzRC7A%2Bi56a7389U7D8FPy2KMSnLdWeLqCIb%2FkeL4%3D&reserved=0
Answer: Thank you very much for the suggestion. We really appreciate the suggestion of Amrhein et al. (2019) paper. We did use effect sizes and its 95% confidence intervals. However, our preference was to provide the absolute mean differences in the units of the outcomes instead of standardised mean difference (e.g., Cohens' d, or Hedges' g effect sizes). In order to include the reviewer's suggestion and enhance clarity, we have added a standardised mean difference (Cohen's D) in the statistical analysis and Table 2 as follows:
Page 8, lines 244-246: "Furthermore, Cohen's d effect size and its 95% CI have been provided where appropriate. According to Cohen (1988), effect size (ES) values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects (Cohen, 1992)."

Page 11-12, Table 2:
Table 2. Muscle strength, thickness and quality, and functional tests absolute values and change over 8 weeks.
VariablesBaseline8 weeksAdjusted mean changeAdjusted group difference
Mean ± SDMean ± SDMean95% CIΔ%Mean95% CIP-valueCohen's d
(95% CI)
Plantar flexor muscle strength
Isometric, N.m
CTR98.4 ± 21.9102.6±29.14.2-6.2 to 14.63.8 ± 14.9%4.6-2.0 to 11.3.1610.6
(-0.2 to 1.4)
RET108.9 ± 17.6123.7±24.114.7*5.3 to 24.213.6 ± 12.3%
Dynamic at 30º.sec-1, N.m
CTR87.4 ± 28.894.2 ± 19.46.8-2.2 to 15.913.6 ± 25.6%-0.2-5.7 to 5.3.928-0.0
(-0.8 to 0.8)
RET86.6 ± 19.993.2 ± 20.36.7*0.1 to 13.28.8 ± 11.4%
Muscle thickness
Plantar flexors, mm
CTR25.5 ± 2.825.1 ± 2.9-0.5-1.5 to 0.6-1.7 ± 6.7%1.30.5 to 2.1.0021.2
(0.5 to 2.3)
RET25.4 ± 2.427.6 ± 3.22.1*0.8 to 3.48.4 ± 7.6%
Gastrocnemius, mm
CTR13.0 ± 1.812.9 ± 2.5-0.1-0.7 to 0.4-1.7 ± 7.7%0.60.2 to 0.9.0061.1
(0.2 to 1.9)
RET12.8 ± 2.213.7 ± 2.50.9*0.4 to 1.47.3 ± 6.5%
Soleus, mm
CTR12.5 ± 3.312.2 ± 3.1-0.3-1.1 to 0.4-1.5 ± 9.9%0.80.2 to 1.3.0081.0
(0.3 to 2.0)
RET12.6 ± 2.113.8 ± 2.71.2*0.3 to 2.19.6 ± 10.2%
MQST
Isometric, N.m.mm-1
CTR3.9 ± 0.84.1 ± 1.00.2-0.2 to 0.65.9 ± 16.1%0.00.0-0.3 to 0.3.9400.0

(-0.8 to 0.8)
RET4.3 ± 0.64.5 ± 0.80.2-0.1 to 0.65.1 ± 12.2%
Dynamic, N.m.mm-1
CTR3.4 ± 1.13.8 ± 0.80.3-0.1 to 0.716.4 ± 30.0%-0.2-0.4 to 0.1.129-0.5
(-1.2 to 0.4)
RET3.4 ± 0.73.4 ± 0.70.0-0.3 to 0.30.8 ± 13.4%
Subcutaneous fat
Plantar flexors, mm
CTR6.8 ± 3.16.8 ± 2.7-0.1-0.5 to 0.31.2 ± 11.9%-0.1-0.4 to 0.3.7370.0
(-0.8 to 0.8)
RET6.5 ± 2.76.3 ± 2.5-0.1-0.8 to 0.50.0 ± 14.1%
MQEI
Plantar flexors, a.u.
CTR20.4 ± 4.723.0 ± 5.22.6*0.0 to 5.115.1 ± 20.8%-0.5-2.0 to 0.9.468-0.2
(-1.0 to 0.6)
RET19.2 ± 3.821.0 ± 4.31.8-0.1 to 3.710.2 ± 15.3%
Gastrocnemius, a.u.
CTR28.1 ± 7.632.6 ± 6.34.5*1.6 to 7.420.9 ± 26.1%-1.7-3.7 to 0.3.0880.6
(-1.4 to 0.2)
RET27.4 ± 5.928.8 ± 6.41.3-1.8 to 4.56.0 ± 18.2%
Soleus, a.u.
CTR12.7 ± 5.613.3 ± 7.30.6-1.7 to 2.96.5 ± 27.2%0.8-0.9 to 2.5.3650.4
(-0.4 to 1.2)
RET11.0 ± 4.313.2 ± 4.82.2-0.4 to 4.828.4 ± 43.6%
Functional tests
Stair climb, sec
CTR8.7 ± 1.38.9 ± 1.10.2-0.3 to 0.62.4 ± 7.6%-0.4-0.7 to -0.1.021-1.3
(-2.1 to -0.4)
RET9.5 ± 2.88.7 ± 2.1-0.8*-1.2 to -0.3-6.1 ± 10.1%
6-m usual walk, sec
CTR4.05 ± 0.214.11 ± 0.230.07-0.02 to 0.161.7 ± 3.5%-0.05-0.17 to 0.08.465-0.8
(-2.7 to -0.8)
RET3.97 ± 0.413.95 ± 0.55-0.21-0.28 to 0.24-0.5 ± 10.0%
TUG, sec
CTR6.24 ± 0.946.41 ± 0.920.16-0.48 to 0.163.0 ± 8.2%-0.1-0.4 to 0.2.403-0.4
(-1.3 to 0.3)
RET6.58 ± 1.276.43 ± 1.17-0.14-0.65 to 0.36-1.5 ± 11.6%
*, Within-groups statistical difference compared to baseline, P<.05. TUG, Timed-up and go test. Cohen's d values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects

After careful reflection, we are following Amrhein et al. (2019) as our discussion did not just rely on P-values, but also extended to 95% CIs (e.g., discussion regarding muscle strength results) and minimal difference needed to be considered real, following the reviewer's suggestion.

Second, the authors need to do a better job highlighting the difference between change scores for the two groups, with an analysis of covariance (dependent variable = posttest, independent variable = group, covariate = pretest) being the preferred approach for pretest-posttest-control group designs. I recommend the authors consult the short review by Bland and Altman (BMJ 2011; 342 doi: https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1 136%2Fbmj.d561&data=02%7C01%7Cplopezda%40our.ecu.edu.au%7C994848 99c5ba44b5982108d7f579daee%7C9bcb323d7fa345e7a36f6d9cfdbcc272%7C1%7C0 %7C637247776028988378&sdata=kp9AU1FafZy%2BxasHzoHPlPD10BdStbZ7Iv XQVX64xBE%3D&reserved=0). Discussion about within-group differences in the absence of an interaction should be avoided.
Answer: Thank you very much for the comment. The original sentence may have been confusing, as a result, the new sentences aim to describe clearly that we used an ANCOVA:
Page 8, lines 236-241: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests or Chi-square tests, where appropriate. After data showed normality and homogeneity

(P> 0.05), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes."

In regard to the point raised about within-group differences and its discussion in the absence of main effects, we found this confusing as may contradict the previous reviewer suggestion (i.e., "The authors' reliance on p values and NHST is not in line with current recommendations"). Furthermore, the reason to provide a discussion on muscle strength was to ensure intervention consistency. It is important to note that the absence of muscle strength improvements is sometimes seen as "a poor resistance training intervention" in the scientific community. Thus, we thought it would be interesting to provide reasons for this and avoid misinterpretations regarding the protocols design.

Finally, coefficient of variation is not an appropriate method for reporting test-retest reliability statistics. Rather, the intraclass correlation coefficient, standard error of measurement, and minimal difference needed to be real should be utilized, as reviewed by Weir (2005; DOI: 10.1519/15184.1). The paper would be greatly enhanced by inclusion of these metrics rather than the coefficient of variation, particularly if the authors can report the number of participants that showed change scores which exceeded the minimal difference needed to be considered real.

Answer: We really appreciate the comment. Intraclass correlation coefficient (ICC) and 'minimal difference needed to be considered real' were provided for the outcomes throughout the text, and coefficient of variation removed. The ICC was high in all outcomes (≥0.93), and the 'minimal difference needed to be considered real' allowed us to identify participants changes' above these values. Furthermore, we used * to denote participants above the 'minimal difference needed to be considered real' values in Figures 2 and 3. Please, see below:

Figure 2. Absolute change in peak torque at 30º.sec-1 (A), isometric peak torque (B) and plantar flexors muscle thickness (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real.

Figure 3. Absolute change in plantar flexors MQEI (A), dynamic MQST (B) and stair climb test (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real for plantar flexors MQEI and stair climb test.

3) I would recommend that the authors explore correcting their echo intensity values for subcutaneous tissue thickness using the equation created by Young et al. (2015, DOI:10.1002/mus.24656). To do so, the authors would need to calculate and report subcutaneous thickness. If the authors do not feel comfortable replacing their traditional values with the corrected values, inclusion of both would bolster the manuscript and make for interesting discussion. There are several excellent echo intensity papers showing that the interpretation of data is greatly affected by subcutaneous thickness correction.

Examples:
Stock et al. (2018, DOI: 10.1016/j.ultrasmedbio.2018.04.009)
Ryan et al. (2016, DOI: 10.1139/apnm-2016-0238)
Oranchuk et al. (2020, DOI: 10.1139/apnm-2019-0601)

Answer: That's a very interesting point! We really appreciate the comment and the studies suggested but we would like to share some thoughts about this specific suggestion.

The study of Young et al. (2015) was a very interesting paper which helps the field to move forward in the investigation of muscle echo intensity. However, we understand that some issues may preclude us to use their equation in our sample:

1) Although the sample size of Young et al. (2015) study was "thirty-one participants (14 men, 17 women) between ages 20 and 61 years", the women's group age ranges from 20 to 29 yrs. Thus, we understand that the women-specific Young's equation for medial gastrocnemius intramuscular fat (i.e., y= [0.239 * (40 * subcutaneous fat thickness) + raw echo intensity] + 4.221) would not be adequate to our sample which

comprised of older women (66.3±5.8 yrs);

2) Although the direction of Young's equation coefficients makes sense (i.e., more subcutaneous fat thickness, brighter would be the echoes after equation correction), the equation could provide values unlikely to be true in our sample, even if corrected by subcutaneous fat tissue. Considering that older adults present less muscle mass because of the sarcopenia process and thereby, decreased fluid storage caused by a lower glycogen-to-muscle area, it is expected that smaller amounts of fluids would be stored (Fernández-Elías et al., 2015, PMID: 25911631) hence affecting muscle echo intensity values (Taniguchi et al., 2017, PMID: 28755131). Thus, young and older women may present different muscular characteristics, which was not accounted for in the Young et al. (2015) study and its formula;

3) The ultrasound device, setup and data acquisition method from our and Young et al. (2015) study were completely different. For example, we used a 38-mm, 9.0 MHz linear array probe, while they utilised a 47-mm multifrequency linear transducer (8-12 MHz). The setup of the images acquisition was 6.0 mm of image depth and 90-dB general gain in our study, while 40 mm and 58-dB were used in Young et al. (2015) study, respectively. Finally, we used an Aloka ultrasound device (Philips, Japan), while Young et al. (2015) have used a LOGIQ e (GE Healthcare, UK). This issue was addressed at page 18, lines 432-435: "Secondly, the variation in MQEI values between ultrasound devices makes it difficult to compare different studies. For example, we used a different ultrasound device, setup and data acquisition method to Young et al. (2015). Thus, the design of an imaging phantom in the future may help to adjust MQEI values of various imaging devices.". Thus, these differences are likely to produce completely different results in muscle echo intensity and precludes comparison between studies;

Thus, we would like to kindly ask the possibility of maintaining the raw echo intensity given the aforementioned. Furthermore, we have reported the subcutaneous fat thickness values to ensure that MQEI results were truly unaltered in our study. Changes were provided as follows:

Pages 6-7, lines 181-186: "Although changes in muscle echo intensity were likely affected by subcutaneous fat thickness (Young et al., 2015), correction equations have not been tested for older populations. Thus, subcutaneous fat thickness values were determined and expressed by the distance between the skin–muscle interface and the superior border of the muscle's aponeurosis using the line tool (Stock et al., 2018). The ICC for subcutaneous fat thickness was 0.97 (standard error mean= 0.1 mm)."

Table 2:

Subcutaneous fat	Baseline	8 weeks	Adjusted mean change			Adjusted group difference		
Plantar flexors, mm	Mean ± SD	Mean ± SD	Mean	95% CI	Δ%	Mean	95% CI	P-value	Cohen's d (95% CI)
CTR	6.8 ± 3.1	6.8 ± 2.7	-0.1	-0.5 to 0.3	1.2 ± 11.9%	-0.1	-0.4 to 0.3	737	0.0 (-0.8 to 0.8)
RET	6.5 ± 2.7	6.3 ± 2.5	-0.1	-0.8 to 0.5	0.0 ± 14.1%				

Page 13, lines 294-295: "Likewise, changes in plantar flexors subcutaneous fat thickness were not observed (P= .737; Table 2)."

4) The authors have made several inferences to the notion that echo intensity reflects only intramuscular adipocyte accumulation. However, the role of fibrous tissue should not be completely discounted. There has also been discussion in the literature that other factors may be at play. I ask that the authors refrain from suggesting that echo intensity only reflects intramuscular adipocyte infiltration, as more research is needed to determine if other factors are at play.

Answer: The reviewer is right, and we agree. Changes were made accordingly at:

Page 3, lines 68-72: "The term muscle quality per se as described by Correa-de-Araujo et al. (2017) refers to two specific measures, intramuscular adipose and fibrous tissue assessment (or non-contractile tissue; e.g., ultrasound-derived muscle echo-intensity (MQEI)), and the relative force production per unit of muscle mass (e.g., expressed as a ratio of peak torque and muscle size; often called muscle specific-tension (MQST))."

Page 14, lines 330-333: "Regarding muscle composition, although intramuscular lipid stores play a role in providing energy substrates during exercise (Pan et al., 1997), its accumulation in conjunction with increases in fibrous tissue within the muscle are elevated in older adults as a result of reduced oxidative capacity (Nakagawa et al., 2007).".

MINOR COMMENTS:

1) Did the authors acquire the gastrocnemius and soleus images together? This is not clear. Also, it is unclear if the gastrocnemius image encompassed both the medial and lateral head. Work by Young et al. (2015, DOI:10.1002/mus.24656) only included the medial gastrocnemius. Please further clarify so that future readers can replicate this study's methods.

Answer: The reviewer is right. We did not specify this important information. We have provided further changes to make it clearer within the methods section:

Page 6, lines 157-160: "Participants rested in the supine position with the lower limbs extended and relaxed for 10 min (Lopez et al., 2018). Similar to a previous study (Stephensen et al., 2014), transverse images of the right medial gastrocnemius and soleus were acquired."

Page 6, lines 162-163: "Three images of the right medial gastrocnemius and soleus were taken together and exported to a personal computer for further analysis, performed by the same investigator.".

2) Please check that the writing throughout the manuscript is in past tense. The study is now over, so the language should not be in present or future tense. For example, "hypothesize" should be "hypothesized" and so on.

Answer: Thank you very much for this comment. We double-checked all the sentences and changes were done throughout the text.

3) The results of the present study are similar to those reported by Mota et al. (2017, DOI: 10.1088/1361-6579/aa791a), but that work was carried out in young boys. As such, the authors should consider if age and/or sex may play a role in these types of studies.

Answer: Unfortunately, after careful consideration of Mota et al. (2017) work, we did not find a way to incorporate a discussion regarding the reviewer's points. The study of Mota et al. (2017) investigated associations between MQEI and MQST in young boys. Our results were related to MQEI and MQST and functional capacity in older women. We tried multiple ways of integrating results from the Mota et al. (2017) study without providing a tangent from our rationale that would likely confuse the reader, yet nothing was forthcoming. Furthermore, we designed an RCT to investigate resistance training effects in a clinical population and outcomes related to functional capacity, different than Mota et al. (2017). We hope that the reviewer understands our rationale and subsequent decision.

4) The abstract could use revision. I have two suggestions. First, demographics of the participants should be included. Second, the first finding within the Results of the abstract should be the lack of change for echo intensity. The change in the stair climb test does not seem to be the main finding of the study; therefore, it should not be discussed first.

Answer: Thank you for the suggestion. Changes were provided within the abstract as follows:

Page 2, lines 31-32: "Twenty-four older women (66.3 ± 5.8 yrs; 69.0 ± 3.0 kg; 25.3 ± 1.4 kg.m-2) were recruited to the study."

Lines 36-40: "After 8 weeks, both MQEI and MQST did not respond to the intervention. Furthermore, significant changes in stair climb performance (P<0.05) were not associated with plantar flexor-derived muscle quality (P>0.05). Finally, significant gains in muscle hypertrophy were observed in the RET group (P<0.01), while muscle strength failed to change significantly (P>0.05).".

5) Line 49: Change / to "and" // 6) Line 53: I recommend adding "physical" prior to the word deterioration. // 7) Lines 81-84: Please revise, as the sentence beginning with "However" is wordy and difficult to follow.

Answer: All suggestions were amended accordingly. Thank you.

8) Lines 105-106: This sentence states that gait speed was part of the study's inclusion/exclusion. Was this assessed prior to enrollment? How did the authors make this determination?

Answer: The reviewer is right; we did not provide the information about this specific exclusion criteria. Changes were made as follows:

Page 4, lines 118-119: "Prior to official enrolment in the study, participants completed a 6 m gait test. Participants were excluded if they had an average speed <1.2 m.s-1.".

9) Line 204: The authors state that their ANOVA was adjusted for baseline values. Is this synonymous with an ANCOVA?

Answer: Yes. We provided changes to make it clearer in our "statistical analysis" section as follows:

Page 8, lines 236-241: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests or Chi-square tests, where appropriate. After data showed normality and homogeneity ($P > 0.05$), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes.".

10) Though randomized control trials are more useful than association-based studies, the findings of the present study are somewhat at odds with results by Mota et al. (2018, DOI:10.1007/s40520-017-0829-1), who reported that echo intensity, but not muscle size, was correlated with muscle performance. The authors may wish to deliberate on this point.

Answer: Thank you for the suggestion. We provided a discussion regarding the results of Mota et al. (2018) as follows:

Page 15, lines 353-365: "Curiously, a non-significant association between MQEI and functional tests' performance was also found in the present study. In contrast to the findings in our present study, Mota et al. (2018) observed a significant negative association between lateral gastrocnemius MQEI levels and a measure related to functional performance (i.e., plantar flexors rate of velocity development). This would indicate that higher levels of MQEI in the lateral gastrocnemius may impair the ability to generate velocity rapidly. However, we did not observe similar effects when evaluating functional capacity itself, possibly given the gastrocnemius portion evaluated (medial vs. lateral) or even the study design (RCT vs. cross-sectional study). Therefore, future studies will be necessary to investigate if changes in functional capacity are mediated by such factors in older adults. Altogether, these results suggest that resistance training improves functional capacity regardless of MQEI adaptations; alternatively, given the lack of changes in MQEI or measures associated with MQEI, it was not possible to observe the translation of this outcome to a better functional capacity."

11) I commend the authors for including information about their a priori power analysis. However, I recommend that the authors include the effect sizes utilized, rather than the change they expected to observe.

Answer: Thank you for the comment. We provided that information as follows:

Page 8, lines 229-235: "The sample size estimate was based on projected changes in muscle quality as measured by MQEI (Wilhelm et al., 2014b). To achieve 80% power at an alpha level of 0.05 (two-tailed), 11 participants per group would be required to detect a mean difference of -4.8 a.u. (standard deviation of 4.0; or an effect size of 1.3) in MQEI. For the secondary outcomes, a sample of 22 participants had sufficient power to detect changes of 0.3 sec in stair climb test (standard deviation of 0.3; or an effect size of 0.5) (Galvão & Taaffe, 2005), and 0.3 sec in TUG test (standard deviation of 0.4; or an effect size of 0.4) (Radaelli et al., 2019)."

Reviewer #2:
Answer: We would like to thank you for the time and effort you gave providing us with constructive comments throughout the manuscript.

Lines 50 and 62: is muscle quality an underlying parameter of the musculoskeletal system? In line 50 it seems like the authors are trying to separate the two, whereas in line 62 like they are parts of a whole? Please revise to provide clarity.

Answer: The reviewer is right. We changed the sentence to make it clearer as follows:

Page 3, lines 60-63: "Among the musculoskeletal system improvements, the benefits on muscle quality have been considered an important target of exercise given its association with functional capacity (Pinto et al., 2014; Fragala et al., 2015; Rech et al., 2014; Wilhelm et al., 2014a; Lopez et al., 2017) […]"

Please check the manuscript for wording about the calf and stick to one for consistency. (e.g. Line 79: "plantar-flexor"; Line 85: "plantar flexor"; Line 86: "calf").
Answer: Thank you for the comment. The manuscript was double-checked accordingly and changes for consistency were provided throughout the text.

Line 139: do you have the average day number following the intervention when post measures were assessed? I think this value would be more useful than the range of days.
Answer: Thank you for the comment. We provided the average number of hours following the intervention. Furthermore, we identified an apparent typo. The post-intervention evaluations were conducted between 3 to 7 days after the final training sessions. Changes were provided at:
Page 6, lines 152-153: "The post-intervention evaluations were performed 78 h (standard deviation of ±10 h) after the completion of the final training session (range: 3 to 7 days)."

Please check the spacing (e.g. "2 min" line 167; "~50min" line 120) throughout the document. There are multiple places where the above issue was spotted.
Answer: Thank you for the comment. We double-checked accordingly and changes were provided throughout the text.

Line 170: why did you use the peak torque values and not the mean values?
Answer: The procedures involving muscle strength evaluation are very common in the literature. In our study, we have cited some manuscripts adopting the same procedure (Radaelli et al., 2019, PMID: 29730331; Radaelli et al., 2014, PMID: 24414336). The intent of assessing muscle strength is to obtain its maximal value whilst avoiding the learning effects regarding the test. Furthermore, participants are rarely familiarised with the isokinetic dynamometer procedures, and even with a familiarisation, variation in strength levels are likely to occur. Moreover, using the mean values would have reduced the muscle strength levels at baseline and further increased the difference pre to post-intervention. Thus, we chose to consider the highest peak torque value in further analysis for consistency and to avoid overestimation of the muscle strength gains attributed to our intervention.

Functional capacity tests: Do you think ceiling effects and your participants baseline health values contributed to not observing changes?
Answer: That's a very interesting point. We are not sure about a ceiling effect as our participants experienced improvements on the functional capacity tests (Cohen's d= -0.4 to -1.3, Table 2). Moreover, the IPAQ levels of our sample may also indicate that physical adaptations are likely to occur. Thus, we are more inclined to believe that both low baseline values and the intervention duration itself precluded the observation of significant differences in our RCT. For clarity within the text, we made changes as follows:
Page 16, lines 377-383: "Furthermore, those previous studies (Galvão & Taaffe, 2005; Radaelli et al., 2019) were longer than the present study (i.e., 12 and 20 weeks) which may indicate that at least 12 weeks would be necessary to observe such changes in 6-m usual walk and TUG test. In hindsight, considering that our sample was mostly participants with moderate to lower levels of physical activity who were untrained in resistance exercise, physical adaptations were likely observable but may have required a larger sample size or a longer period of intervention.".

Line 204: ANOVA or ANCOVA with adjusted values? Overall, the stats section was not very clear. Why were effect sizes not used for a training intervention? A number of times the authors cited that the lack of data makes it difficult to compare to other studies, but effect sizes can easily be compared.
Answer: Thank you very much for the comment. ANCOVA was used in our analysis. We updated the "Statistical Analysis" section to make it clearer. Actually, we used effect sizes and its 95% confidence intervals in the units of the outcomes (see in 'adjusted mean difference' and 'adjusted group difference' in Table 2). In accordance with both reviewers' comments we provided Cohens' d in Table 2 and changes within the statistical analysis section as follows:
Page 8, lines 236-246: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests

or Chi-square tests, where appropriate. After data showed normality and homogeneity (P> 0.05), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes. Bonferroni post hoc procedure for multiple comparisons was conducted if the interaction time x group or main effect for time was significant to locate the source of the significant differences. Tests were two-tailed with statistical significance set at an alpha level of .05. Furthermore, Cohen's d effect size and its 95% CI have been provided where appropriate. According to Cohen (1988), effect size (ES) values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects (Cohen, 1992).".

Regarding comparisons with other studies, unfortunately, just the Cohens' d values are unlikely to help us in this issue. The main issue is the lack of studies evaluating plantar flexors muscle quality (i.e., muscle hypertrophy, echo intensity or specific tension). Therefore, as we proposed a unique RCT evaluating this specific measure, limitations regarding current literature preclude a more comprehensive discussion regarding plantar flexors MQEI.

Figure 1: please change "no lost of follow-up" to "no lost during follow-up"
Answer: Thank you for the comment. Figure 1 was changed accordingly.

Table 1: were there any differences in the raw IPAQ scores between groups?
Answer: No differences were found between groups on scores or Kcal.wk-1 derived from IPAQ (P= .634 and .801). The Kcal.wk-1 values observed were 1,201 ± 398 for RET and 1,144 ± 398 for CTR group.
We provided this information as follows:
Table 1. Baseline characteristics of the participants.
VariablesOverall
(n= 24)RET
(n= 12)CTR
(n= 12)
Age, mean ± SD, yr66.3 ± 5.867.1 ± 6.363.3 ± 5.6
Weight, mean ± SD, kg69.0 ± 3.069.8 ± 2.768.0 ± 3.4
Height, mean ± SD, cm165.0 ± 3.5165 ± 3.6166 ± 3.8
BMI, mean ± SD, kg.m-225.3 ± 1.425.6 ± 1.324.9 ± 0.7
IPAQ score
High, N (%)3 (12.5)1 (8.3)2 (16.6)
Moderate, N (%)8 (33.3)5 (41.6)3 (24.9)
Low, N (%)13 (54.1)6 (50)7 (58.3)
IPAQ, International Physical Activity Questionnaire. Differences were not observed between groups (P> .05).
Page 9, lines 251-253: "The groups were balanced at baseline (Table 1) and participants did not present any differences between IPAQ scores (P= .634), as well as no comorbidities before and during the study."

It is not clear if the isometric, dynamic, or both contractions were used for MQ calculations.
Answer: We regret to have not included such important information before. Thank you very much for this comment. A subsection was provided for MQST as follows:
Page 7, lines 204-209: "2.4.3. Specific tension
MQST was calculated relative to the MVIC and plantar flexor muscle strength at 30º.sec-1. Thus, the isometric MQST was determined by the ratio between MVIC and plantar flexor muscle thickness values, while the ratio between plantar flexor muscle

strength at 30º.sec-1 and plantar flexor muscle thickness values were used to determine the dynamic MQST. Both ratios were expressed as N.m.mm-1."
Figure 2: not all of the figures have dashed lines, if you add them they should be added to each of the figures.
Answer: The reviewer is right. Thank you for the comment. Although dashed lines were introduced in all figures, some of them were over the x-axis line (e.g., Figure 3, panel B) and hard to see. We have subsequently increased the thickness of the dashed lines, improving their visibility in all figures.

Figure 2. Absolute change in peak torque at 30º.sec-1 (A), isometric peak torque (B) and plantar flexors muscle thickness (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real.

Figure 3. Absolute change in plantar flexors MQEI (A), dynamic MQST (B) and stair climb test (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real for plantar flexors MQEI and stair climb test.
Although the others suggest it is an important finding, do the authors think that the change in strength is clinically relevant or meaningful. When comparing CIs there is no difference, and the mean change for the CTR is actually greater than that of the RET group.
Answer: This is a very interesting point-of-view raised by the reviewer. We agree that changes in muscle strength might not be an important finding in clinical trials. Nonetheless, it is interesting that changes in this outcome are commonly used to determine resistance training effectiveness. One of the reasons to describe this outcome is to show that even without superiority for plantar flexor muscle strength, the RET group were the only ones to improve functional capacity. Furthermore, we also thought it important to avoid any misinterpretation of our training program given the lack of changes in muscle strength. Thus, we decided to reorganise our discussion section, which now flows as such:
1) Main findings; 2) MQEI; 3) MQEI vs. Functional capacity; 4) Functional capacity; 5) Muscle hypertrophy/ MQST; 6) Muscle strength; 7) Strengths and limitations.

How is it plausible that the calf muscles may experience larger hypetrophic adaptations in comparison to strength during short-term interventions? The points made above about 1) fiber type and 2) minimal stimulus contradict this. Could you expand this thought process further?
Answer: Honestly, we were as surprised as the reviewer regarding this result. The hypothesis stated at page 4, lines 105-108: "Since muscle size adaptation is more prominent in long-term interventions, we hypothesised that resistance training would promote significant improvements in muscle strength, muscle quality and functional capacity, but not muscle thickness." indicated that our expectation was to observe the opposite. After careful consideration of this result, we reflect that our baseline values might have moderated the gains induced by resistance training. Furthermore, the number of participants above the "minimal difference needed to be considered true" also supports this finding.
Since the '80s, the classic work from Sale (1988, PMID: 3057313) and its Figure 11 became very popular in resistance training science (cited more than 1,500 times) and the 'neural vs. muscular adaptations to strength training' idea disseminated for more than 30 years. We agree with this model. However, we also understand that Sale's resistance training adaptations model is unlikely to be extended to all muscles and populations. The plantar flexor muscles are poorly investigated in current literature, even more so in older adults. Thus, we reorganised the paragraph regarding muscle hypertrophy discussion to make clear our assumption about the lower baseline levels as follows:
Pages 16-17, lines 384-405: "The significant increase in plantar flexors muscle thickness following a short-term resistance training program was unexpected in the present study. Significant increases of ~2.0 mm was found on plantar flexors muscle thickness following 8 weeks of resistance training (n= 9 above the minimal difference

needed to be considered real). The reasons for this may be related to the baseline values of our sample. Although the gastrocnemius and soleus muscles are primarily comprised of slow-twitch fibres (i.e., type I) at ~60 and 80% (Gollnick et al., 1974), respectively, the participants presented a gastrocnemius muscle thickness of ~13.0mm, similar to sarcopenic older adults in the studies of Kuyumcu et al. (2016) and Wang et al. (2018) (15.0 and 13.7mm, respectively). Thus, even with an attenuated response in these muscles, given the lower hypertrophic potential compared to fast-twitch fibres (Fry, 2004), the resistance training intervention was likely to induce a significant increase in plantar flexors muscle thickness moderated by the low baseline values. This might explain partially the positive effect on this outcome after a short period of intervention. Furthermore, the significant increase in muscle thickness was also associated with the lack of changes in MQST. Following a short-term resistance training program in untrained participants, we would have expected to observe more neural (i.e., muscle strength) than morphological alterations (i.e., muscle hypertrophy) (Sale, 1988). To the contrary, we observed that the plantar flexor muscles of older women may not respond in that way, resulting in a non-significant change in MQST between groups. Thus, dissimilar to results observed in quadriceps femoris muscles (Pinto et al., 2014; Radaelli et al., 2014), it is suggestive that plantar flexor muscles are more likely to present changes in muscle size rather than strength gains following a short-term intervention, particularly when the participants present at baseline with reduced muscle mass levels. Future studies are necessary to elucidate further mechanisms.".

Professor Christiaan Leeuwenburgh
Editor-in-Chief
Experimental Gerontology

May 27th, 2020

Dear Professor Christiaan Leeuwenburgh,

We thank you for the opportunity to revise and resubmit our manuscript. Please find attached the last version of the paper entitled "Effects of an 8-week resistance training intervention on plantar flexor muscle quality and functional capacity in older women: a randomised controlled trial", by Pedro Lopez, Brendan James Crosby, Bruna Patrícia Robetti, Douglas Jean Preussler Turella, Thaís Andréia Schepa Weber, Morgana Lima de Oliveira and Anderson Rech, which has been revised according to the comments and suggestion of the reviewers.

A point-by-point response to the reviewers' comments is provided below and following is a summary of the main changes made in the paper:
* Abstract was changed following Reviewer #1 comments;
* Suggestions for consistency were incorporated and highlighted throughout the text accordingly to Reviewer #1 and #2;
* Settings and participants section were amended with Reviewer #1 comment;
* Ultrasound procedures and analysis section was clarified following Reviewer #1 comments;
* ICC and 'minimal difference needed to be considered real' were provided for the main outcomes as suggested by Reviewer #1;
* The "Specific tension" section was provided following Reviewer #2 comment;
* Standardised mean difference effects were provided for sample size calculations and results, following Reviewer #1 and #2 comments;
* Figure 1, 2 and 3 were amended following Reviewer #1 and #2 comments;
* Discussion section was reorganised to clarify the results importance in our study following Reviewer #2 comments;
* All questions were answered and subsequently referenced when necessary, and when changes were not provided, further explanation were provided.

Sincerely,

Pedro Lopez, MSc
Exercise Medicine Research Institute
Edith Cowan University, AUSTRALIA

Reviewer #1

Skeletal muscle echo intensity and the notion of muscle quality are novel research concepts that are being reported in the literature on a more frequent basis. Several previous studies have reported significant reductions in echo intensity in response to resistance training in both younger and older adults. However, many previous investigators have focused their attention on the quadriceps femoris muscles. As such, the purpose of the present study was to examine the effects of a eight week resistance training intervention on echo intensity, muscle quality (specific tension), and functional outcomes in older adults, with a secondary purpose of examining the associations among change scores. I believe that this study is of interest to the research community and is an excellent fit for Experimental Gerontology. However, I have several significant concerns that need to be addressed prior to acceptance. I have provided my comments below in chronological order separated into major and minor comments. Intervention studies such as the present one are very challenging to carry out, so I commend the authors for their efforts.

Answer: We really appreciate the positive and constructive comments from the reviewer. Thank you very much.

MAJOR COMMENTS:

1) One of my main concerns about the study was whether the training volume for the calf musculature was sufficient to elicit meaningful changes in muscle morphology. If the authors' emphasis was on the calf musculature, why were exercises for the upper limbs included with so few exercises that directly targeted the tested muscles? This needs to be considered and deliberated in the discussion section.

Answer: Thank you for the comment. This question has been very important during and after the conception of our study. We would like to share some thoughts on this.

First, previous studies have prescribed one (e.g., leg extension) (Radaelli et al., 2019, PMID: 29730331) or two resistance exercises (e.g., leg press and leg extension) (Radaelli et al., 2014, PMID: 24414336; Wilhelm et al., 2014, PMID: 25449853) for quadriceps femoris adaptations. These studies reported significant changes (~5 to 20%; P< 0.05) on muscle quality using single or multiple sets (Radaelli et al., 2019, PMID: 29730331; Radaelli et al., 2014, PMID: 24414336; Wilhelm et al., 2014, PMID: 25449853). Thus, if single sets of one or two resistance exercises were sufficient to stimulate changes in quadriceps femoris echo intensity, a prescription involving standing calf raises, 3 sets of 6-12 repetitions at 8-15RM was hypothesised to be equally capable of providing similar changes in calf muscles. However, differences were not observed for 8 weeks, suggesting that calf muscles may need greater local stimulus. Thus, our resistance training prescription design was in accordance with previous literature but produced different results than expected.

We have addressed this issue at:

Pages 3, lines 79-82: "Thus, different exercise modes, or even lower volumes of resistance training (i.e., single sets, or 1-2 resistance exercises) can promote changes to quadriceps derived MQEI by non-contractile tissue reduction as suggested by the authors (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b)."

Page 15, lines 345-349: "Thirdly, although quadriceps femoris may respond to a relatively low exercise stimulus as observed in previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014), the proposed number of exercises, intensity or volume (standing

calf raise, 3 sets of 6-12 repetitions at 8-15RM) might have been an insufficient stimulus for plantar flexor muscles based on this muscle groups level of activity.".

Moreover, during the conception of this trial, we believed that a general resistance training program would be in accordance with exercise guidelines for older adults (Fragala et al., 2019, PMID: 31343601) and thereby, more appropriate than 2-3 calf specific resistance exercises. We were interested in providing all the possible benefits exercise affords older adults, and not just those tested by ourselves.

In order to make it clearer, we provided changes at:

Page 18, lines 440-447: "Finally, our resistance training prescription did not solely target the plantar flexor muscles and may not have provided a stimulus sufficient to improve MQ_{EI} or functional test results. However, previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b) demonstrated that quadriceps femoris MQ_{EI} was likely to present changes in response to single sets of one or two resistance exercises. With this in mind, we targeted multiple muscle groups in accordance with the latest exercise guideline for older adults, focusing on the overall benefits for functional capacity (Fragala et al., 2019) and not exclusively those tested by ourselves."

2) The authors' approach to the statistical analysis could use revision. I have three main qualms.

First, the authors' primary tool for examining both changes between and withing groups should be effect sizes, followed by 95% confidence intervals. The authors' reliance on p values and NHST is not in line with current recommendations. Please see:

https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.nature.com%2Farticles%2Fd41586-019-00857-9&data=02%7C01%7Cplopezda%40our.ecu.edu.au%7C99484899c5ba44b5982108d7f579daee%7C9bcb323d7fa345e7a36f6d9cfdbcc272%7C1%7C0%7C637247776028988378&sdata=p6FzRC7A%2Bi56a7389U7D8FPy2KMSnLdWeLqCIb%2FkeL4%3D&reserved=0

Answer: Thank you very much for the suggestion. We really appreciate the suggestion of Amrhein et al. (2019) paper. We did use effect sizes and its 95% confidence intervals.

However, our preference was to provide the absolute mean differences in the units of the outcomes instead of standardised mean difference (e.g., Cohens' d, or Hedges' g effect sizes). In order to include the reviewer's suggestion and enhance clarity, we have added a standardised mean difference (Cohen's D) in the statistical analysis and Table 2 as follows:

Page 8, lines 244-246: "Furthermore, Cohen's d effect size and its 95% CI have been provided where appropriate. According to Cohen (1988), effect size (ES) values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects (Cohen, 1992)."

Table 2. Muscle strength, thickness and quality, and functional tests absolute values and change over 8 weeks.

Variables	Baseline	8 weeks	Adjusted mean change			Adjusted group difference			
	Mean ± SD	Mean ± SD	Mean	95% CI	Δ%	Mean	95% CI	P-value	Cohen's d (95% CI)
Plantar flexor muscle strength									
Isometric, N.m									
CTR	98.4 ± 21.9	102.6±29.1	4.2	-6.2 to 14.6	3.8 ± 14.9%	4.6	-2.0 to 11.3	.161	0.6 (-0.2 to 1.4)
RET	108.9 ± 17.6	123.7±24.1	14.7*	5.3 to 24.2	13.6 ± 12.3%				
Dynamic at $30°.sec^{-1}$, N.m									
CTR	87.4 ± 28.8	94.2 ± 19.4	6.8	-2.2 to 15.9	13.6 ± 25.6%	-0.2	-5.7 to 5.3	.928	-0.0 (-0.8 to 0.8)
RET	86.6 ± 19.9	93.2 ± 20.3	6.7*	0.1 to 13.2	8.8 ± 11.4%				
Muscle thickness									
Plantar flexors, mm									
CTR	25.5 ± 2.8	25.1 ± 2.9	-0.5	-1.5 to 0.6	-1.7 ± 6.7%	1.3	0.5 to 2.1	**.002**	1.2 (0.5 to 2.3)
RET	25.4 ± 2.4	27.6 ± 3.2	2.1*	0.8 to 3.4	8.4 ± 7.6%				
Gastrocnemius, mm									
CTR	13.0 ± 1.8	12.9 ± 2.5	-0.1	-0.7 to 0.4	-1.7 ± 7.7%	0.6	0.2 to 0.9	**.006**	1.1 (0.2 to 1.9)
RET	12.8 ± 2.2	13.7 ± 2.5	0.9*	0.4 to 1.4	7.3 ± 6.5%				
Soleus, mm									
CTR	12.5 ± 3.3	12.2 ± 3.1	-0.3	-1.1 to 0.4	-1.5 ± 9.9%	0.8	0.2 to 1.3	**.008**	1.0 (0.3 to 2.0)
RET	12.6 ± 2.1	13.8 ± 2.7	1.2*	0.3 to 2.1	9.6 ± 10.2%				
MQ_{ST}									
Isometric, $N.m.mm^{-1}$									
CTR	3.9 ± 0.8	4.1 ± 1.0	0.2	-0.2 to 0.6	5.9 ± 16.1%	0.0	-0.3 to 0.3	.940	0.0 (-0.8 to 0.8)
RET	4.3 ± 0.6	4.5 ± 0.8	0.2	-0.1 to 0.6	5.1 ± 12.2%				
Dynamic, $N.m.mm^{-1}$									
CTR	3.4 ± 1.1	3.8 ± 0.8	0.3	-0.1 to 0.7	16.4 ± 30.0%	-0.2	-0.4 to 0.1	.129	-0.5 (-1.2 to 0.4)
RET	3.4 ± 0.7	3.4 ± 0.7	0.0	-0.3 to 0.3	0.8 ± 13.4%				
Subcutaneous fat									
Plantar flexors, mm									
CTR	6.8 ± 3.1	6.8 ± 2.7	-0.1	-0.5 to 0.3	1.2 ± 11.9%	-0.1	-0.4 to 0.3	.737	0.0 (-0.8 to 0.8)
RET	6.5 ± 2.7	6.3 ± 2.5	-0.1	-0.8 to 0.5	0.0 ± 14.1%				
MQ_{EI}									

	Pre	Post	Change	95% CI	% change	d	95% CI	P	d (95% CI)
<mark>Plantar flexors</mark>, a.u.									
CTR	20.4 ± 4.7	23.0 ± 5.2	2.6*	0.0 to 5.1	15.1 ± 20.8%	-0.5	-2.0 to 0.9	.468	<mark>-0.2 (-1.0 to 0.6)</mark>
RET	19.2 ± 3.8	21.0 ± 4.3	1.8	-0.1 to 3.7	10.2 ± 15.3%				
Gastrocnemius, a.u.									
CTR	28.1 ± 7.6	32.6 ± 6.3	4.5*	1.6 to 7.4	20.9 ± 26.1%	-1.7	-3.7 to 0.3	.088	<mark>0.6 (-1.4 to 0.2)</mark>
RET	27.4 ± 5.9	28.8 ± 6.4	1.3	-1.8 to 4.5	6.0 ± 18.2%				
Soleus, a.u.									
CTR	12.7 ± 5.6	13.3 ± 7.3	0.6	-1.7 to 2.9	6.5 ± 27.2%	0.8	-0.9 to 2.5	.365	<mark>0.4 (-0.4 to 1.2)</mark>
RET	11.0 ± 4.3	13.2 ± 4.8	2.2	-0.4 to 4.8	28.4 ± 43.6%				
Functional tests									
Stair climb, sec									
CTR	8.7 ± 1.3	8.9 ± 1.1	0.2	-0.3 to 0.6	2.4 ± 7.6%	-0.4	-0.7 to -0.1	**.021**	<mark>-1.3 (-2.1 to -0.4)</mark>
RET	9.5 ± 2.8	8.7 ± 2.1	-0.8*	-1.2 to -0.3	-6.1 ± 10.1%				
6-m usual walk, sec									
CTR	4.05 ± 0.21	4.11 ± 0.23	0.07	-0.02 to 0.16	1.7 ± 3.5%	-0.05	-0.17 to 0.08	.465	<mark>-0.8 (-2.7 to -0.8)</mark>
RET	3.97 ± 0.41	3.95 ± 0.55	-0.21	-0.28 to 0.24	-0.5 ± 10.0%				
TUG, sec									
CTR	6.24 ± 0.94	6.41 ± 0.92	0.16	-0.48 to 0.16	3.0 ± 8.2%	-0.1	-0.4 to 0.2	.403	<mark>-0.4 (-1.3 to 0.3)</mark>
RET	6.58 ± 1.27	6.43 ± 1.17	-0.14	-0.65 to 0.36	-1.5 ± 11.6%				

*, Within-groups statistical difference compared to baseline, P<.05. TUG, Timed-up and go test. <mark>Cohen's d values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects</mark>

After careful reflection, we are following Amrhein et al. (2019) as our discussion did not just rely on P-values, but also extended to 95% CIs (e.g., discussion regarding muscle strength results) and *minimal difference needed to be considered real*, following the reviewer's suggestion.

Second, the authors need to do a better job highlighting the difference between change scores for the two groups, with an analysis of covariance (dependent variable = posttest, independent variable = group, covariate = pretest) being the preferred approach for pretest-posttest-control group designs. I recommend the authors consult the short review by Bland and Altman (BMJ 2011; 342 doi: https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1136%2Fbmj.d561&data=02%7C01%7Cplopezda%40our.ecu.edu.au%7C99484899c5ba44b5982108d7f579daee%7C9bcb323d7fa345e7a36f6d9cfdbcc272%7C1%7C0%7C637247776028988378&sdata=kp9AU1FafZy%2BxasHzoHPlPD10BdStbZ7IvXQVX64xBE%3D&reserved=0). Discussion about within-group differences in the absence of an interaction should be avoided.

Answer: Thank you very much for the comment. The original sentence may have been confusing, as a result, the new sentences aim to describe clearly that we used an ANCOVA:

Page 8, lines 236-241: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests or Chi-square tests, where appropriate. After data showed normality and homogeneity ($P > 0.05$), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes."

In regard to the point raised about within-group differences and its discussion in the absence of main effects, we found this confusing as may contradict the previous reviewer suggestion (i.e., "The authors' reliance on p values and NHST is not in line with current recommendations"). Furthermore, the reason to provide a discussion on muscle strength was to ensure intervention consistency. It is important to note that the absence of muscle strength improvements is sometimes seen as "a poor resistance training

intervention" in the scientific community. Thus, we thought it would be interesting to provide reasons for this and avoid misinterpretations regarding the protocols design.

Finally, coefficient of variation is not an appropriate method for reporting test-retest reliability statistics. Rather, the intraclass correlation coefficient, standard error of measurement, and minimal difference needed to be real should be utilized, as reviewed by Weir (2005; DOI: 10.1519/15184.1). The paper would be greatly enhanced by inclusion of these metrics rather than the coefficient of variation, particularly if the authors can report the number of participants that showed change scores which exceeded the minimal difference needed to be considered real.

Answer: We really appreciate the comment. Intraclass correlation coefficient (ICC) and 'minimal difference needed to be considered real' were provided for the outcomes throughout the text, and coefficient of variation removed. The ICC was high in all outcomes ($\geq$0.93), and the 'minimal difference needed to be considered real' allowed us to identify participants changes' above these values. Furthermore, we used * to denote participants above the 'minimal difference needed to be considered real' values in Figures 2 and 3. Please, see below:
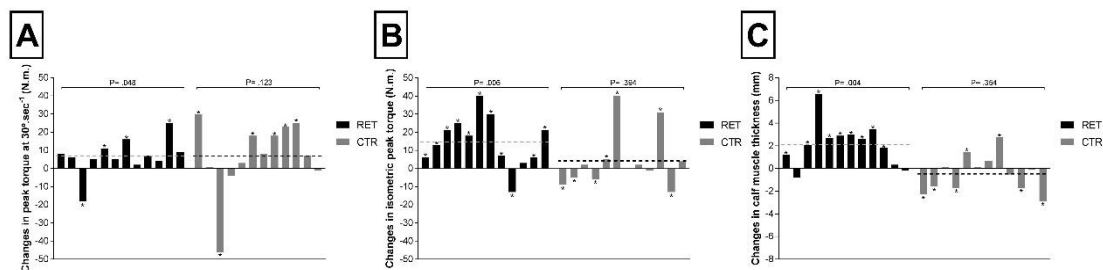

Figure 2. Absolute change in peak torque at 30º.sec-1 (A), isometric peak torque (B) and plantar flexors muscle thickness (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real.
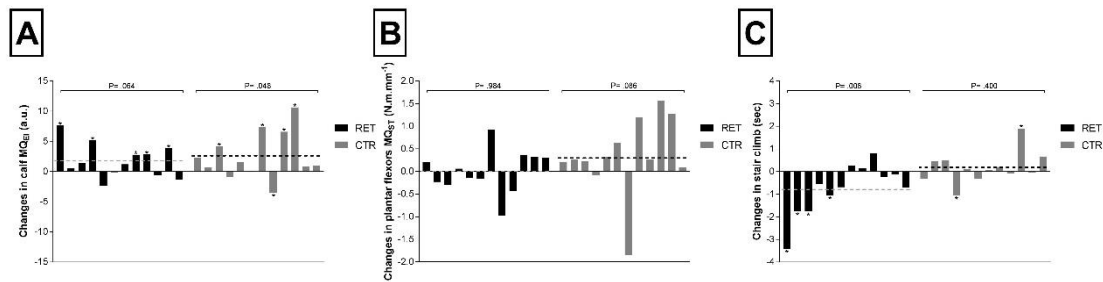
Figure 3. Absolute change in plantar flexors MQ_{EI} (A), dynamic MQ_{ST} (B) and stair climb test (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real for plantar flexors MQEI and stair climb test.

3) I would recommend that the authors explore correcting their echo intensity values for subcutaneous tissue thickness using the equation created by Young et al. (2015, DOI:10.1002/mus.24656). To do so, the authors would need to calculate and report subcutaneous thickness. If the authors do not feel comfortable replacing their traditional values with the corrected values, inclusion of both would bolster the manuscript and make for interesting discussion. There are several excellent echo intensity papers showing that the interpretation of data is greatly affected by subcutaneous thickness correction.

Examples:

Stock et al. (2018, DOI: 10.1016/j.ultrasmedbio.2018.04.009)

Ryan et al. (2016, DOI: 10.1139/apnm-2016-0238)

Oranchuk et al. (2020, DOI: 10.1139/apnm-2019-0601)

Answer: That's a very interesting point! We really appreciate the comment and the studies suggested but we would like to share some thoughts about this specific suggestion.

The study of Young et al. (2015) was a very interesting paper which helps the field to move forward in the investigation of muscle echo intensity. However, we understand that some issues may preclude us to use their equation in our sample:

1) Although the sample size of Young et al. (2015) study was "thirty-one participants (14 men, 17 women) between ages 20 and 61 years", the women's group age ranges from 20 to 29 yrs. Thus, we understand that the women-specific Young's equation for medial gastrocnemius intramuscular fat (i.e., y= [0.239 * (40 * subcutaneous fat thickness) + raw echo intensity] + 4.221) would not be adequate to our sample which comprised of older women (66.3 ± 5.8 yrs);

2) Although the direction of Young's equation coefficients makes sense (i.e., more subcutaneous fat thickness, brighter would be the echoes after equation correction), the equation could provide values unlikely to be true in our sample, even if corrected by subcutaneous fat tissue. Considering that older adults present less muscle mass because of the sarcopenia process and thereby, decreased fluid storage caused by a lower glycogen-to-muscle area, it is expected that smaller amounts of fluids would be stored (Fernández-Elías et al., 2015, PMID: 25911631) hence affecting muscle echo intensity values (Taniguchi et al., 2017, PMID: 28755131). Thus, young and older women may present different muscular characteristics, which was not accounted for in the Young et al. (2015) study and its formula;

3) The ultrasound device, setup and data acquisition method from our and Young et al. (2015) study were completely different. For example, we used a 38-mm, 9.0 MHz linear array probe, while they utilised a 47-mm multifrequency linear transducer (8-12 MHz). The setup of the images acquisition was 6.0 mm of image depth and 90-dB general gain in our study, while 40 mm and 58-dB were used in Young et al. (2015) study, respectively. Finally, we used an Aloka ultrasound device (Philips, Japan), while Young et al. (2015) have used a LOGIQ e (GE Healthcare, UK). This issue was addressed at page 18, lines 432-435: "Secondly, the variation in MQEI values between ultrasound devices makes it difficult to compare different studies. For example, we used a different ultrasound device, setup and data acquisition method to Young et al. (2015). Thus, the design of an imaging phantom in the future may help to adjust MQEI values of various imaging devices.". Thus, these differences are likely to produce completely different results in muscle echo intensity and precludes comparison between studies;

Thus, we would like to kindly ask the possibility of maintaining the raw echo intensity given the aforementioned. Furthermore, we have reported the subcutaneous fat thickness values to ensure that MQ_{EI} results were truly unaltered in our study. Changes were provided as follows:

Pages 6-7, lines 181-186: "Although changes in muscle echo intensity were likely affected by subcutaneous fat thickness (Young et al., 2015), correction equations have not been tested for older populations. Thus, subcutaneous fat thickness values were determined and expressed by the distance between the skin–muscle interface and the superior border of the muscle's aponeurosis using the line tool (Stock et al., 2018). The ICC for subcutaneous fat thickness was 0.97 (standard error mean= 0.1 mm)."

Table 2:

Subcutaneous fat	Baseline	8 weeks	Adjusted mean change			Adjusted group difference			
Plantar flexors, mm	Mean ± SD	Mean ± SD	Mean	95% CI	Δ%	Mean	95% CI	P-value	Cohen's d (95% CI)
CTR	6.8 ± 3.1	6.8 ± 2.7	-0.1	-0.5 to 0.3	1.2 ± 11.9%	-0.1	-0.4 to 0.3	.737	0.0 (-0.8 to 0.8)
RET	6.5 ± 2.7	6.3 ± 2.5	-0.1	-0.8 to 0.5	0.0 ± 14.1%				

Page 13, lines 294-295: "Likewise, changes in plantar flexors subcutaneous fat thickness were not observed (P= .737; Table 2)."

4) The authors have made several inferences to the notion that echo intensity reflects only intramuscular adipocyte accumulation. However, the role of fibrous tissue should not be completely discounted. There has also been discussion in the literature that other factors may be at play. I ask that the authors refrain from suggesting that echo intensity only reflects intramuscular adipocyte infiltration, as more research is needed to determine if other factors are at play.

Answer: The reviewer is right, and we agree. Changes were made accordingly at:

Page 3, lines 68-72: "The term muscle quality per se as described by Correa-de-Araujo et al. (2017) refers to two specific measures, intramuscular adipose and fibrous tissue assessment (or non-contractile tissue; e.g., ultrasound-derived muscle echo-intensity (MQ_{EI})), and the relative force production per unit of muscle mass (e.g., expressed as a ratio of peak torque and muscle size; often called muscle specific-tension (MQ_{ST}))."

Page 14, lines 330-333: "Regarding muscle composition, although intramuscular lipid stores play a role in providing energy substrates during exercise (Pan et al., 1997), its accumulation in conjunction with increases in fibrous tissue within the muscle are elevated in older adults as a result of reduced oxidative capacity (Nakagawa et al., 2007).".

MINOR COMMENTS:

1) Did the authors acquire the gastrocnemius and soleus images together? This is not clear. Also, it is unclear if the gastrocnemius image encompassed both the medial and lateral head. Work by Young et al. (2015, DOI:10.1002/mus.24656) only included the medial gastrocnemius. Please further clarify so that future readers can replicate this study's methods.

Answer: The reviewer is right. We did not specify this important information. We have provided further changes to make it clearer within the methods section:

Page 6, lines 157-160: "Participants rested in the supine position with the lower limbs extended and relaxed for 10 min (Lopez et al., 2018). Similar to a previous study (Stephensen et al., 2014), transverse images of the right medial gastrocnemius and soleus were acquired."

Page 6, lines 162-163: "Three images of the right medial gastrocnemius and soleus were taken together and exported to a personal computer for further analysis, performed by the same investigator.".

2) Please check that the writing throughout the manuscript is in past tense. The study is now over, so the language should not be in present or future tense. For example, "hypothesize" should be "hypothesized" and so on.

Answer: Thank you very much for this comment. We double-checked all the sentences and changes were done throughout the text.

3) The results of the present study are similar to those reported by Mota et al. (2017, DOI: 10.1088/1361-6579/aa791a), but that work was carried out in young boys. As such, the authors should consider if age and/or sex may play a role in these types of studies.

Answer: Unfortunately, after careful consideration of Mota et al. (2017) work, we did not find a way to incorporate a discussion regarding the reviewer's points. The study of Mota et al. (2017) investigated associations between MQ_{EI} and MQ_{ST} in young boys. Our results were related to MQ_{EI} and MQ_{ST} and functional capacity in older women. We tried

multiple ways of integrating results from the Mota et al. (2017) study without providing a tangent from our rationale that would likely confuse the reader, yet nothing was forthcoming. Furthermore, we designed an RCT to investigate resistance training effects in a clinical population and outcomes related to functional capacity, different than Mota et al. (2017). We hope that the reviewer understands our rationale and subsequent decision.

4) The abstract could use revision. I have two suggestions. First, demographics of the participants should be included. Second, the first finding within the Results of the abstract should be the lack of change for echo intensity. The change in the stair climb test does not seem to be the main finding of the study; therefore, it should not be discussed first.

Answer: Thank you for the suggestion. Changes were provided within the abstract as follows:

Page 2, lines 31-32: "Twenty-four older women (66.3 ± 5.8 yrs; 69.0 ± 3.0 kg; 25.3 ± 1.4 kg.m-2) were recruited to the study."

Lines 36-40: "After 8 weeks, both MQEI and MQST did not respond to the intervention. Furthermore, significant changes in stair climb performance ($P<0.05$) were not associated with plantar flexor-derived muscle quality ($P>0.05$). Finally, significant gains in muscle hypertrophy were observed in the RET group ($P<0.01$), while muscle strength failed to change significantly ($P>0.05$).".

5) Line 49: Change / to "and" // 6) Line 53: I recommend adding "physical" prior to the word deterioration. // 7) Lines 81-84: Please revise, as the sentence beginning with "However" is wordy and difficult to follow.

Answer: All suggestions were amended accordingly. Thank you.

8) Lines 105-106: This sentence states that gait speed was part of the study's inclusion/exclusion. Was this assessed prior to enrollment? How did the authors make this determination?

Answer: The reviewer is right; we did not provide the information about this specific exclusion criteria. Changes were made as follows:

Page 4, lines 118-119: "Prior to official enrolment in the study, participants completed a 6 m gait test. Participants were excluded if they had an average speed <1.2 $m.s^{-1}$.".

9) Line 204: The authors state that their ANOVA was adjusted for baseline values. Is this synonymous with an ANCOVA?

Answer: Yes. We provided changes to make it clearer in our "statistical analysis" section as follows:

Page 8, lines 236-241: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests or Chi-square tests, where appropriate. After data showed normality and homogeneity (P> 0.05), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes.".

10) Though randomized control trials are more useful than association-based studies, the findings of the present study are somewhat at odds with results by Mota et al. (2018, DOI:10.1007/s40520-017-0829-1), who reported that echo intensity, but not muscle size, was correlated with muscle performance. The authors may wish to deliberate on this point.

Answer: Thank you for the suggestion. We provided a discussion regarding the results of Mota et al. (2018) as follows:

Page 15, lines 353-365: "Curiously, a non-significant association between MQ_{EI} and functional tests' performance was also found in the present study. In contrast to the findings in our present study, Mota et al. (2018) observed a significant negative association between lateral gastrocnemius MQ_{EI} levels and a measure related to

functional performance (i.e., plantar flexors rate of velocity development). This would indicate that higher levels of MQ_{EI} in the lateral gastrocnemius may impair the ability to generate velocity rapidly. However, we did not observe similar effects when evaluating functional capacity itself, possibly given the gastrocnemius portion evaluated (medial vs. lateral) or even the study design (RCT vs. cross-sectional study). Therefore, future studies will be necessary to investigate if changes in functional capacity are mediated by such factors in older adults. Altogether, these results suggest that resistance training improves functional capacity regardless of MQ_{EI} adaptations; alternatively, given the lack of changes in MQ_{EI} or measures associated with MQ_{EI}, it was not possible to observe the translation of this outcome to a better functional capacity."

11) I commend the authors for including information about their a priori power analysis. However, I recommend that the authors include the effect sizes utilized, rather than the change they expected to observe.

Answer: Thank you for the comment. We provided that information as follows:

Page 8, lines 229-235: "The sample size estimate was based on projected changes in muscle quality as measured by MQ_{EI} (Wilhelm et al., 2014b). To achieve 80% power at an alpha level of 0.05 (two-tailed), 11 participants per group would be required to detect a mean difference of -4.8 a.u. (standard deviation of 4.0; or an effect size of 1.3) in MQ_{EI}. For the secondary outcomes, a sample of 22 participants had sufficient power to detect changes of 0.3 sec in stair climb test (standard deviation of 0.3; or an effect size of 0.5) (Galvão & Taaffe, 2005), and 0.3 sec in TUG test (standard deviation of 0.4; or an effect size of 0.4) (Radaelli et al., 2019)."

Answer: We would like to thank you for the time and effort you gave providing us with constructive comments throughout the manuscript.

Lines 50 and 62: is muscle quality an underlying parameter of the musculoskeletal system? In line 50 it seems like the authors are trying to separate the two, whereas in line 62 like they are parts of a whole? Please revise to provide clarity.

Answer: The reviewer is right. We changed the sentence to make it clearer as follows:

Page 3, lines 60-63: "Among the musculoskeletal system improvements, the benefits on muscle quality have been considered an important target of exercise given its association with functional capacity (Pinto et al., 2014; Fragala et al., 2015; Rech et al., 2014; Wilhelm et al., 2014a; Lopez et al., 2017) […]"

Please check the manuscript for wording about the calf and stick to one for consistency. (e.g. Line 79: "plantar-flexor"; Line 85: "plantar flexor"; Line 86: "calf").

Answer: Thank you for the comment. The manuscript was double-checked accordingly and changes for consistency were provided throughout the text.

Line 139: do you have the average day number following the intervention when post measures were assessed? I think this value would be more useful than the range of days.

Answer: Thank you for the comment. We provided the average number of hours following the intervention. Furthermore, we identified an apparent typo. The post-intervention evaluations were conducted between 3 to 7 days after the final training sessions. Changes were provided at:

Page 6, lines 152-153: "The post-intervention evaluations were performed 78 h (standard deviation of ±10 h) after the completion of the final training session (range: 3 to 7 days)."

Please check the spacing (e.g. "2 min" line 167; "~50min" line 120) throughout the document. There are multiple places where the above issue was spotted.

Answer: Thank you for the comment. We double-checked accordingly and changes were provided throughout the text.

Line 170: why did you use the peak torque values and not the mean values?

Answer: The procedures involving muscle strength evaluation are very common in the literature. In our study, we have cited some manuscripts adopting the same procedure (Radaelli et al., 2019, PMID: 29730331; Radaelli et al., 2014, PMID: 24414336). The intent of assessing muscle strength is to obtain its maximal value whilst avoiding the learning effects regarding the test. Furthermore, participants are rarely familiarised with the isokinetic dynamometer procedures, and even with a familiarisation, variation in strength levels are likely to occur. Moreover, using the mean values would have reduced the muscle strength levels at baseline and further increased the difference pre to post-intervention. Thus, we chose to consider the highest peak torque value in further analysis for consistency and to avoid overestimation of the muscle strength gains attributed to our intervention.

Functional capacity tests: Do you think ceiling effects and your participants baseline health values contributed to not observing changes?

Answer: That's a very interesting point. We are not sure about a ceiling effect as our participants experienced improvements on the functional capacity tests (Cohen's d= -0.4 to -1.3, Table 2). Moreover, the IPAQ levels of our sample may also indicate that physical adaptations are likely to occur. Thus, we are more inclined to believe that both low baseline values and the intervention duration itself precluded the observation of significant differences in our RCT. For clarity within the text, we made changes as follows:

Page 16, lines 377-383: "Furthermore, those previous studies (Galvão & Taaffe, 2005; Radaelli et al., 2019) were longer than the present study (i.e., 12 and 20 weeks) which may indicate that at least 12 weeks would be necessary to observe such changes in 6-m usual walk and TUG test. In hindsight, considering that our sample was mostly

participants with moderate to lower levels of physical activity who were untrained in resistance exercise, physical adaptations were likely observable but may have required a larger sample size or a longer period of intervention.".

Line 204: ANOVA or ANCOVA with adjusted values? Overall, the stats section was not very clear. Why were effect sizes not used for a training intervention? A number of times the authors cited that the lack of data makes it difficult to compare to other studies, but effect sizes can easily be compared.

Answer: Thank you very much for the comment. ANCOVA was used in our analysis. We updated the "Statistical Analysis" section to make it clearer. Actually, we used effect sizes and its 95% confidence intervals in the units of the outcomes (see in 'adjusted mean difference' and 'adjusted group difference' in Table 2). In accordance with both reviewers' comments we provided Cohens' d in Table 2 and changes within the statistical analysis section as follows:

Page 8, lines 236-246: "Normality of the distribution was assessed by the Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-tests or Chi-square tests, where appropriate. After data showed normality and homogeneity (P> 0.05), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline values were used to assess change over time (baseline and 8 weeks) in the primary and secondary outcomes. Bonferroni post hoc procedure for multiple comparisons was conducted if the interaction time x group or main effect for time was significant to locate the source of the significant differences. Tests were two-tailed with statistical significance set at an alpha level of .05. Furthermore, Cohen's d effect size and its 95% CI have been provided where appropriate. According to Cohen (1988), effect size (ES) values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects (Cohen, 1992).".

Regarding comparisons with other studies, unfortunately, just the Cohens' d values are unlikely to help us in this issue. The main issue is the lack of studies evaluating plantar flexors muscle quality (i.e., muscle hypertrophy, echo intensity or specific tension). Therefore, as we proposed a unique RCT evaluating this specific measure, limitations regarding current literature preclude a more comprehensive discussion regarding plantar flexors MQ_{EI}.

Figure 1: please change "no lost of follow-up" to "no lost during follow-up"

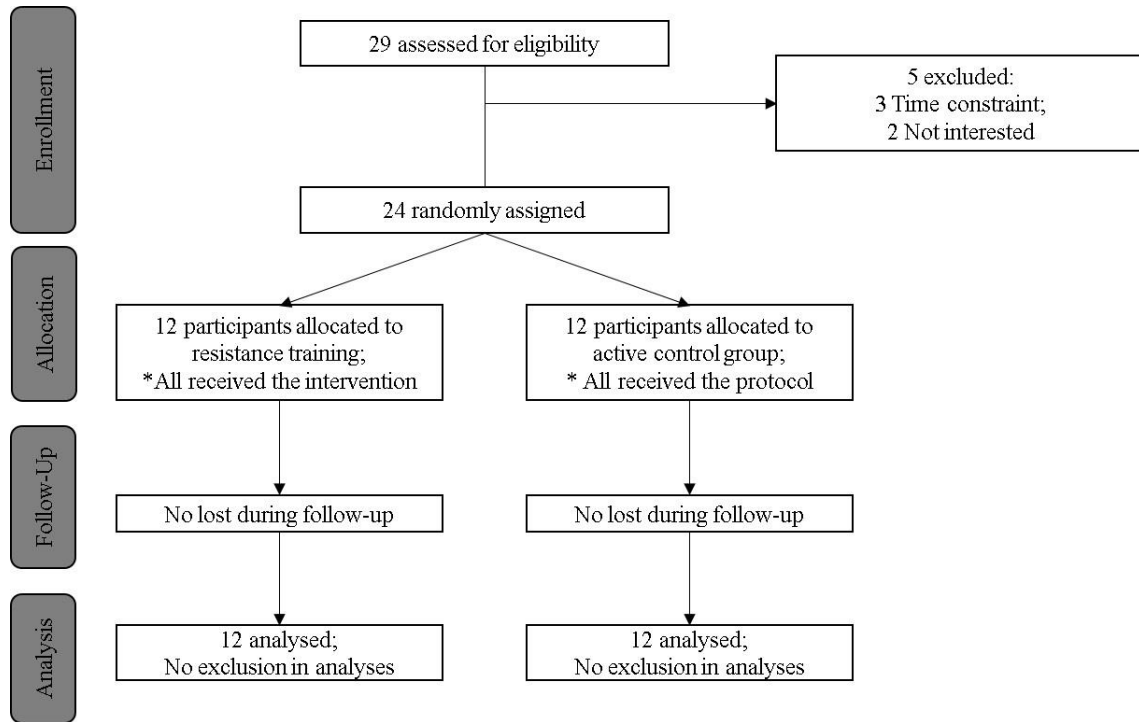**Answer:** Thank you for the comment. Figure 1 was changed accordingly.

Table 1: were there any differences in the raw IPAQ scores between groups?

Answer: No differences were found between groups on scores or Kcal.wk^{-1} derived from IPAQ (P= .634 and .801). The Kcal.wk^{-1} values observed were 1,201 ± 398 for RET and 1,144 ± 398 for CTR group.

We provided this information as follows:

Table 1. Baseline characteristics of the participants.

Variables	Overall (n= 24)	RET (n= 12)	CTR (n= 12)
Age, mean ± SD, yr	66.3 ± 5.8	67.1 ± 6.3	63.3 ± 5.6
Weight, mean ± SD, kg	69.0 ± 3.0	69.8 ± 2.7	68.0 ± 3.4
Height, mean ± SD, cm	165.0 ± 3.5	165 ± 3.6	166 ± 3.8
BMI, mean ± SD, kg.m^{-2}	25.3 ± 1.4	25.6 ± 1.3	24.9 ± 0.7
IPAQ score			
High, N (%)	3 (12.5)	1 (8.3)	2 (16.6)
Moderate, N (%)	8 (33.3)	5 (41.6)	3 (24.9)
Low, N (%)	13 (54.1)	6 (50)	7 (58.3)

IPAQ, International Physical Activity Questionnaire. Differences were not observed between groups (P> .05).

Page 9, lines 251-253: "The groups were balanced at baseline (Table 1) and participants did not present any differences between IPAQ scores (P= .634), as well as no comorbidities before and during the study."

It is not clear if the isometric, dynamic, or both contractions were used for MQ calculations.

Answer: We regret to have not included such important information before. Thank you very much for this comment. A subsection was provided for MQ$_{ST}$ as follows:

Page 7, lines 204-209: "2.4.3. Specific tension

MQ$_{ST}$ was calculated relative to the MVIC and plantar flexor muscle strength at 30º.sec-1. Thus, the isometric MQ$_{ST}$ was determined by the ratio between MVIC and plantar flexor muscle thickness values, while the ratio between plantar flexor muscle

strength at 30º.sec-1 and plantar flexor muscle thickness values were used to determine the dynamic MQ_{ST}. Both ratios were expressed as N.m.mm-1."

Figure 2: not all of the figures have dashed lines, if you add them they should be added to each of the figures.

Answer: The reviewer is right. Thank you for the comment. Although dashed lines were introduced in all figures, some of them were over the x-axis line (e.g., Figure 3, panel B) and hard to see. We have subsequently increased the thickness of the dashed lines, improving their visibility in all figures.

Figure 2. Absolute change in peak torque at 30º.sec-1 (A), isometric peak torque (B) and plantar flexors muscle thickness (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real.
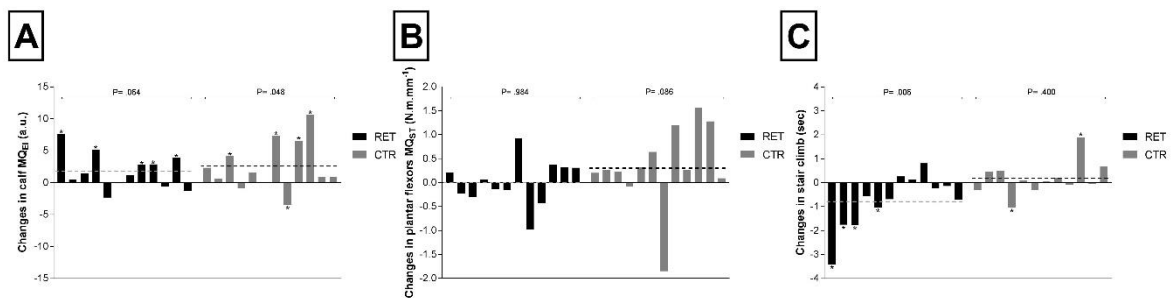

Figure 3. Absolute change in plantar flexors MQ_{EI} (A), dynamic MQ_{ST} (B) and stair climb test (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real for plantar flexors MQ_{EI} and stair climb test.

Although the others suggest it is an important finding, do the authors think that the change in strength is clinically relevant or meaningful. When comparing CIs there is no difference, and the mean change for the CTR is actually greater than that of the RET group.

Answer: This is a very interesting point-of-view raised by the reviewer. We agree that changes in muscle strength might not be an important finding in clinical trials. Nonetheless, it is interesting that changes in this outcome are commonly used to determine resistance training effectiveness. One of the reasons to describe this outcome is to show that even without superiority for plantar flexor muscle strength, the RET group were the only ones to improve functional capacity. Furthermore, we also thought it important to avoid any misinterpretation of our training program given the lack of changes in muscle strength. Thus, we decided to reorganise our discussion section, which now flows as such:

1) Main findings; 2) MQ_{EI}; 3) MQ_{EI} vs. Functional capacity; 4) Functional capacity; 5) Muscle hypertrophy/ MQ_{ST}; 6) Muscle strength; 7) Strengths and limitations.

How is it plausible that the calf muscles may experience larger hypetrophic adaptations in comparison to strength during short-term interventions? The points made above about 1) fiber type and 2) minimal stimulus contradict this. Could you expand this thought process further?

Answer: Honestly, we were as surprised as the reviewer regarding this result. The hypothesis stated at page 4, lines 105-108: "Since muscle size adaptation is more prominent in long-term interventions, we hypothesised that resistance training would promote significant improvements in muscle strength, muscle quality and functional capacity, but not muscle thickness." indicated that our expectation was to observe the opposite. After careful consideration of this result, we reflect that our baseline values might have moderated the gains induced by resistance training. Furthermore, the number of participants above the "minimal difference needed to be considered true" also supports this finding.

Since the '80s, the classic work from Sale (1988, PMID: 3057313) and its Figure 11 became very popular in resistance training science (cited more than 1,500 times) and

the 'neural vs. muscular adaptations to strength training' idea disseminated for more than 30 years. We agree with this model. However, we also understand that Sale's resistance training adaptations model is unlikely to be extended to all muscles and populations. The plantar flexor muscles are poorly investigated in current literature, even more so in older adults. Thus, we reorganised the paragraph regarding muscle hypertrophy discussion to make clear our assumption about the lower baseline levels as follows:

Pages 16-17, lines 384-405: "The significant increase in plantar flexors muscle thickness following a short-term resistance training program was unexpected in the present study. Significant increases of ~2.0 mm was found on plantar flexors muscle thickness following 8 weeks of resistance training (n= 9 above the minimal difference needed to be considered real). The reasons for this may be related to the baseline values of our sample. Although the gastrocnemius and soleus muscles are primarily comprised of slow-twitch fibres (i.e., type I) at ~60 and 80% (Gollnick et al., 1974), respectively, the participants presented a gastrocnemius muscle thickness of ~13.0mm, similar to sarcopenic older adults in the studies of Kuyumcu et al. (2016) and Wang et al. (2018) (15.0 and 13.7mm, respectively). Thus, even with an attenuated response in these muscles, given the lower hypertrophic potential compared to fast-twitch fibres (Fry, 2004), the resistance training intervention was likely to induce a significant increase in plantar flexors muscle thickness moderated by the low baseline values. This might explain partially the positive effect on this outcome after a short period of intervention. Furthermore, the significant increase in muscle thickness was also associated with the lack of changes in MQST. Following a short-term resistance training program in untrained participants, we would have expected to observe more neural (i.e., muscle strength) than morphological alterations (i.e., muscle hypertrophy) (Sale, 1988). To the contrary, we observed that the plantar flexor muscles of older women may not respond in that way, resulting in a non-significant change in MQST between groups. Thus, dissimilar to results observed in quadriceps femoris muscles (Pinto et al., 2014; Radaelli et al., 2014), it is suggestive that plantar flexor muscles are more likely to present changes in muscle size rather than strength gains following a short-term intervention, particularly when the participants present at baseline with reduced muscle mass levels. Future studies are necessary to elucidate further mechanisms.".

Highlights

- Muscle quality has been considered an important target of resistance exercise.

- Muscle quality has been associated with functional capacity.

- No investigation has been conducted for calf derived muscle quality.

- Resistance training provides significant benefits in stair climb performance.

- Gains in stair climb performance were not associated with calf muscle quality.

1 **EFFECTS OF AN 8-WEEK RESISTANCE TRAINING INTERVENTION ON**
2 **PLANTAR FLEXOR MUSCLE QUALITY AND FUNCTIONAL CAPACITY IN**
3 **OLDER WOMEN: A RANDOMISED CONTROLLED TRIAL**

4 Pedro Lopez[1*], Brendan James Crosby[1], Bruna Patrícia Robetti[2], Douglas Jean Preussler

5 Turella[3], Thaís Andréia Schepa Weber[2], Morgana Lima de Oliveira[2], Anderson Rech[2]

6 1 - Exercise Medicine Research Institute, Edith Cowan University, Perth, Western Australia,

7 Australia

8 2 – Universidade de Caxias do Sul, Caxias do Sul, Rio Grande do Sul, Brazil

9 3 – Centro Clínico UCS, Universidade de Caxias do Sul, Caxias do Sul, Rio Grande do Sul,

10 Brazil

11

12 * - Corresponding author

13 Pedro Lopez, MSc

14 Exercise Medicine Research Institute

15 Edith Cowan University

16 270 Joondalup Drive, Joondalup WA 6027

17 AUSTRALIA

18 T +61 416463228

19 Email: p.lopezda@our.ecu.edu.au

20

21

22

23

24

25

26

ABSTRACT

The present study examined 8 weeks of resistance training and its effects on muscle quality measures, plantar flexor muscle strength, muscle thickness and functional capacity in older women. Moreover, we tested if changes in muscle quality were associated with functional capacity. Twenty-four older women (66.3 ± 5.8 yrs; 69.0 ± 3.0 kg; 25.3 ± 1.4 kg.m^{-2}) were recruited to the study. After completion of the baseline assessment, participants were randomly assigned to either the resistance training (RET, n= 12) or an active control group (CTR, n= 12). Muscle quality was evaluated through muscle echo intensity (MQ$_{EI}$) and specific tension (MQ$_{ST}$). Muscle thickness, unilateral plantar flexor muscle strength and functional tests were evaluated at baseline and after the training period. After 8 weeks, both MQ$_{EI}$ and MQ$_{ST}$ did not respond to the intervention. Furthermore, significant changes in stair climb performance ($P<0.05$) were not associated with plantar flexor-derived muscle quality ($P>0.05$). Finally, significant gains in muscle hypertrophy were observed in the RET group ($P<0.01$), while muscle strength failed to change significantly ($P>0.05$). In conclusion, a resistance training program provided significant benefits in the stair climb test, unrelated to plantar flexor-derived muscle quality measures as previously demonstrated in quadriceps femoris.

Keywords: Aging; Resistance training; Muscle quality; Muscle echo intensity; Physical function.

1. INTRODUCTION

There is a consensus in the literature for utilising resistance training to benefit older adults (ACSM, 2009; Fragala et al., 2019). Resistance training can significantly improve functional capacity and an aging neuromuscular system; increasing the threshold for disabilities whilst reducing subsequent falls and mortality risk in older adults (Fragala et al., 2019). Among the musculoskeletal system improvements, the benefits on muscle quality have been considered an important target of exercise given its association with functional capacity (Pinto et al., 2014; Fragala et al., 2015; Rech et al., 2014; Wilhelm et al., 2014a; Lopez et al., 2017) and the natural physical deterioration that occurs over a lifespan (Lynch et al., 1999; Arts et al., 2009). Thus, strategies to counter and mitigate the aging-related decline in muscle quality could help maintain physical function in older adults as preconized by the *World Health Organization* concept of healthy aging (Beard et al., 2016).

The term muscle quality *per se* as described by Correa-de-Araujo et al. (2017) refers to two specific measures, intramuscular adipose and fibrous tissue assessment (or non-contractile tissue; e.g., ultrasound-derived muscle echo-intensity (MQ_{EI})), and the relative force production per unit of muscle mass (e.g., expressed as a ratio of peak torque and muscle size; often called muscle specific-tension (MQ_{ST})). Although a vague term, the measurement of muscle quality provides insights into age-related musculoskeletal deterioration and potential strategies to counteract the changes in muscle metabolism, structure and function. However, when restricted to resistance training studies in older adults, muscle quality measures are often derived from quadriceps femoris muscles. For example, improvements in MQ_{EI} range from ~5 to 20% in older adults following resistance training (Radaelli et al., 2014), combined resistance and aerobic training (Wilhelm et al., 2014b) and high-velocity resistance training (Radaelli et al., 2019). Thus, different exercise modes, or even lower volumes of resistance training (i.e., single sets, or 1-2 resistance exercises) can promote changes to quadriceps derived MQ_{EI} by non-contractile tissue reduction as suggested by the authors (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b). Likewise, MQ_{ST} improvements in response to resistance training are also consistent, ranging from ~15-22% following 6 and 13 weeks of training (Pinto et al., 2014; Radaelli et al., 2013;2014). The reason for marked improvements in MQ_{ST} appears to be mediated by non-hypertrophic-related factors, with muscle strength and hypertrophy affected differently by short-periods of intervention (i.e., priority for neural rather than morphological adaptations) (Sale, 1988). Thus, although positively affected by resistance training, which may improve physical reserve (Buchner & deLateur, 1991), it is unknown if lower-extremity

89 muscles, other than the quadriceps femoris, respond similarly to a resistance training program
90 or even alter functional capacity in older adults.

91 The ankle plantar flexor muscles play a major role in gait and stair climbing,
92 independently predicting the variation on these functional tests in older women as
93 demonstrated by the study of Suzuki et al. (2001). However, the number of interventional
94 studies focusing on the plantar-flexor muscles is limited, a few have investigated the effects on
95 muscle strength or power (Capodaglio et al., 2005; Gavin et al., 2019), though none have
96 reported on morphological changes (muscle thickness or muscle quality). Consequently, it is
97 unknown if resistance training can induce changes in plantar flexor muscle quality; and if that,
98 in turn, translates into a better functional capacity. Exploring plantar flexor derived muscle
99 quality may elucidate if resistance training-induced changes on functional capacity could be
100 explained by different muscle group features other than quadriceps femoris. This information
101 may help to design more effective resistance training programs aimed at improving functional
102 capacity in older adults. Therefore, the purpose of the present study was to examine 8 weeks
103 of progressive resistance training and its effects on muscle quality measures, muscle strength
104 and muscle thickness derived from plantar flexor muscles, and functional capacity in older
105 women. Furthermore, we tested if changes in muscle quality were related to functional
106 capacity. Since muscle size adaptation is more prominent in long-term interventions, we
107 hypothesised that resistance training would promote significant improvements in muscle
108 strength, muscle quality and functional capacity, but not muscle thickness. We also
109 hypothesised that changes in muscle quality were associated with changes in functional
110 capacity tests.

111

112 **2. MATERIALS AND METHODS**

113 *2.1. Settings and participants*

114 Twenty-four older women were recruited to the study between July 2017 and August
115 2018, by oral invitation or advertisement in local media (e.g., social networks). Inclusion
116 criteria included an age ≥ 60 years, BMI <30 kg.m^{-2}. The exclusion criterion was verified by
117 questionnaire and included uncontrolled hypertension, a musculoskeletal impairment that
118 restricted physical exercise, and participation in any regular physical exercise program within
119 the 6 months prior to enrolment in the study. Prior to official enrolment in the study,
120 participants completed a 6 m gait test. Participants were excluded if they had an average speed

121 $<1.2 \text{ m.s}^{-1}$. The study procedures were approved by the Research Ethics Committee of Caxias

122 do Sul University (approval number 2687471), with all procedures conducted following the

123 ethical principles of the World Medical Association (Declaration of Helsinki), with all patients

124 providing their written informed consent to participate.

2.2. Study design and random assignment

126 This study was a two-armed, prospective RCT. After the completion of the baseline

127 assessment, participants were randomly assigned to one of the two arms: resistance training

128 (RET) or active control group (CTR). The randomisation sequence was computer-generated

129 according to randomised block design, stratified by age and with 1:1 ratio, and allocation

130 concealment was done by an independent researcher, blinded to the details of the study.

2.3. Resistance training program

132 Participants in the RET group undertook a resistance training program twice per week

133 for 8 weeks. Sessions were conducted in small groups (3 to 5) of participants under the direct

134 supervision of an exercise physiologist. The sessions were ~50 min in duration, commencing

135 with a 5 min warm-up period consisting of low-intensity aerobic exercise on the treadmill (11-

136 to 13- point Borg Scale). The resistance exercises included chest press, squat, lateral pulldown,

137 standing calf raise, lateral raise, bilateral knee extension, abdominal crunches, and knee flexion.

138 The programme was designed to progress in loading from 15- to 8- repetition maximum (RM)

139 for 3 sets per exercise. The number of repetitions was designed to progress from 12 to 6

140 repetitions allowing a safety margin of 2-3 repetitions in each set. All exercises were performed

141 utilising a 2 sec concentric and eccentric phase, with a 45-60 sec rest period between sets.

142 For the CTR group, participants undertook low-intensity joint mobilisation and static

143 stretches prior to 20 mins low-intensity aerobic exercise twice a week. The sessions target

144 intensity was 11- to 13- points on the Borg Scale. During the study, the participants were

145 encouraged to maintain customary activity levels and dietary patterns. Self-reported physical

146 activity was assessed using the International Physical Activity Questionnaires (IPAQ).

2.4. Primary and secondary endpoints

148 Study endpoints were assessed at baseline and after the 8-week intervention. The

149 primary study endpoint was plantar flexor derived MQ_{EI} used as a measure of muscle quality

150 (Lopez et al., 2017; Rech et al., 2014; Wilhelm et al., 2014a). Secondary endpoints were MQ_{ST},

151 plantar flexor muscle strength, muscle and subcutaneous thickness, and functional capacity

tests. Except for the functional capacity tests, all the evaluations were administered by assessors blinded to group assignment. The post-intervention evaluations were performed 78 h (standard deviation of ±10 h) after the completion of the final training session (range: 3 to 7 days).

2.4.1. Ultrasound procedures and analysis

Plantar flexor muscles B-mode ultrasound images were obtained with a 38-mm, 9.0 MHz linear array probe (image depth: 6.0 mm, 90-dB general gain, time-gain compensation in the neutral position) using an ultrasound device (Philips Aloka, Tokyo, Japan). Participants rested in the supine position with the lower limbs extended and relaxed for 10 min (Lopez et al., 2018). Similar to a previous study (Stephensen et al., 2014), transverse images of the right medial gastrocnemius and soleus were acquired. The measurement was taken at 33% of the distance between the lateral condyle of the femur and the lateral malleolus (Stephensen et al., 2014). Three images of the right medial gastrocnemius and soleus were taken together and exported to a personal computer for further analysis, performed by the same investigator. Image analyses were performed using ImageJ 1.42q software (National Institutes of Health, Bethesda, MD, USA).

Muscle thickness was determined by the distance between the adipose tissue to muscle interface for gastrocnemius, and as the bone to muscle interface for soleus. Image analyses were performed in the ImageJ 1.42q software using the line tool. Plantar flexor muscle thickness was obtained as the sum of the gastrocnemius (intraclass correlation coefficient (ICC)= 0.97; standard error mean= 0.27 mm) and soleus muscle thickness (ICC= 0.97; standard error mean= 0.28 mm). The intraclass correlation coefficient (ICC) for the plantar flexor muscle thickness was 0.97 (standard error mean= 0.27 mm; minimal difference needed to be considered real= 0.75 mm). Furthermore, MQ_{EI} was determined by the regions of interest for the gastrocnemius and soleus, which include the selection of as much muscle as possible while avoiding bone and surrounding fascia. The mean echo intensity was determined using a standard grey-scale histogram function and expressed as a value between 0 (black) and 255 (white) for each muscle in arbitrary units (a.u.). Plantar flexor MQ_{EI} was determined from the average echo intensity values from gastrocnemius and soleus muscle portions. The MQ_{EI} ICC for gastrocnemius and soleus was 0.95 (standard error mean= 1.1 a.u.) and 0.92, (standard error mean= 0.7 a.u.). The Plantar flexor MQ_{EI} ICC was 0.93 (standard error mean= 0.9 a.u.; minimal difference needed to be considered real= 2.4 a.u.).

183   Although changes in muscle echo intensity were likely affected by subcutaneous fat
184   thickness (Young et al., 2015), correction equations have not been tested for older populations.
185   Thus, subcutaneous fat thickness values were determined and expressed by the distance
186   between the skin–muscle interface and the superior border of the muscle's aponeurosis using
187   the line tool (Stock et al., 2018). The ICC for subcutaneous fat thickness was 0.97 (standard
188   error mean= 0.1 mm).

189   *2.4.2. Isokinetic dynamometer*

190   The maximum voluntary isometric contractions (MVIC) of the right plantar flexor was
191   assessed using an isokinetic dynamometer (Byodex, USA). The joint angle of plantar flexion
192   was 0°, assuming 0° as neutral positioning of the joint. Participants performed a standard
193   familiarization protocol and following a 2 min rest interval were instructed to perform three
194   maximum contractions. There was an interval of 90 sec between each of the three trials. All
195   subjects were encouraged during the test and instructed to perform the contraction "as fast and
196   strong as possible" (Sahaly et al., 2001). The MVIC with the highest peak torque was
197   considered for further analyses. The isokinetic device was calibrated according to the
198   manufacturer's instructions. The ICC for this measure was 0.98 (standard error mean= 1.6 N.m;
199   minimal difference needed to be considered real= 4.6 N.m). Furthermore, a dynamic
200   assessment of plantar flexor muscle strength was also performed at $30°.sec^{-1}$. The range of
201   motion was 0° (assuming 0° as neutral positioning of the joint) to 45°, with a 90s interval given
202   between each of the three trials. A standard familiarization involving submaximal contractions
203   was done before the valid attempts. The highest concentric peak torque value was used for
204   further analyses. The ICC for this measure was 0.95 (standard error mean= 3.2 N.m; minimal
205   difference needed to be considered real= 8.8 N.m).

206   *2.4.3. Specific tension*

207   MQ_{ST} was calculated relative to the MVIC and plantar flexor muscle strength at $30°.sec^{-1}$
208   . Thus, the isometric MQ_{ST} was determined by the ratio between MVIC and plantar flexor
209   muscle thickness values, while the ratio between plantar flexor muscle strength at $30°.sec^{-1}$ and
210   plantar flexor muscle thickness values were used to determine the dynamic MQ_{ST}. Both ratios
211   were expressed as $N.m.mm^{-1}$.

212   *2.4.4. Functional capacity tests*

213 Before the performance of the functional tests, participants were familiarized with the
214 protocols. The stair climb test constituted the time it takes to go up and down a flight of stairs
215 (10 stairs per flight, 20-cm rise per stair) at their usual pace. Three attempts were made, and
216 the average performance of the attempts was used for further analysis. The ICC for stair climbs
217 trials was 0.96 (standard error mean= 0.3 sec; minimal difference needed to be considered real=
218 0.8 sec). In the timed-up and go (TUG) test, subjects were seated in a hard-backed chair (43
219 cm from the floor) with their arms folded across their chest. After a predetermined signal, they
220 were instructed to rise as fast as possible without the aid of their arms and walk in a fast manner
221 for a distance of 3 meters, turn around and return (via the same route) to sit in the chair again.
222 The lowest time of three attempts was considered for further analysis. The ICC for the TUG
223 test was 0.95 (standard error mean= 0.2 sec; minimal difference needed to be considered real=
224 0.5 sec). In the 6-m usual walk, the participants were required to walk 10 meters at a normal
225 pace. The initial and final 2 meters were disregarded due to the acceleration and deceleration
226 periods, respectively. The valid time for the test refers to the intermediate 6 meters. The test
227 was performed three times and the average time to perform it was used to calculate the usual
228 walking speed (Green et al., 2002). The ICC for the 6-m usual walk test was 0.96 (standard
229 error mean= 0.2 sec; minimal difference needed to be considered real= 0.6 sec).

230 ### 2.5. Statistical analysis and sample size calculation

231 The sample size estimate was based on projected changes in muscle quality as measured
232 by MQ_{EI} (Wilhelm et al., 2014b). To achieve 80% power at an alpha level of 0.05 (two-tailed),
233 11 participants per group would be required to detect a mean difference of -4.8 a.u. (standard
234 deviation of 4.0; or an effect size of 1.3) in MQ_{EI}. For the secondary outcomes, a sample of 22
235 participants had sufficient power to detect changes of 0.3 sec in stair climb test (standard
236 deviation of 0.3; or an effect size of 0.5) (Galvão & Taaffe, 2005), and 0.3 sec in TUG test
237 (standard deviation of 0.4; or an effect size of 0.4) (Radaelli et al., 2019). Data was analysed
238 using SPSS v.22 (IBM Corp., NY, USA). Normality of the distribution was assessed by the
239 Shapiro-Wilk test. Between-group differences at baseline were assessed using independent T-
240 tests or Chi-square tests, where appropriate. After data showed normality and homogeneity (P>
241 0.05), two-way repeated-measures analysis of covariance (ANCOVA) adjusted for baseline
242 values were used to assess change over time (baseline and 8 weeks) in the primary and
243 secondary outcomes. Bonferroni *post hoc* procedure for multiple comparisons was conducted
244 if the interaction time x group or main effect for time was significant to locate the source of the
245 significant differences. Tests were two-tailed with statistical significance set at an alpha level

of .05. <mark>Furthermore, Cohen's d effect size and its 95% CI have been provided where appropriate. According to Cohen (1988), effect size (ES) values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8 indicate medium, and values ≥0.8 indicate large effects (Cohen, 1992).</mark>

3. RESULTS

3.1. Participants characteristics

Twenty-four (82.7%) out of the twenty-nine screened participants were recruited to the study (Figure 1). The most common reasons for non-participation were time constraints and disagreement with the study protocol. <mark>The groups were balanced at baseline (Table 1) and participants did not present any differences between IPAQ scores (P= .634), as well as no comorbidities before and during the study.</mark> The RET and CTR groups attended 84% (161 of 192 sessions), and 76% (145 of 192 sessions), respectively. Participants in the RET group presented an attendance of 13.9 ± 1.5 sessions, while attendance in the CTR group was of 11.3 ± 3.4 sessions.
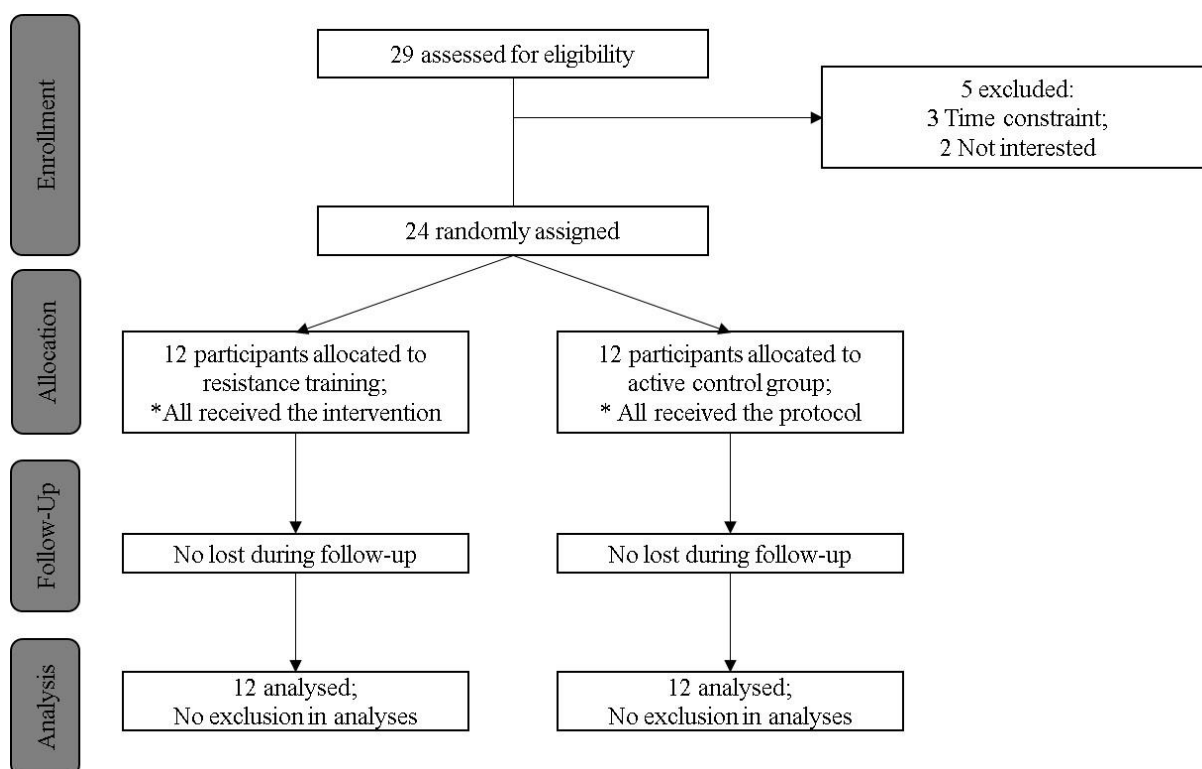
Figure 1. Participant flowchart.

263 **Table 1.** Baseline characteristics of the participants.

Variables	Overall (n= 24)	RET (n= 12)	CTR (n= 12)
Age, mean ± SD, yr	66.3 ± 5.8	67.1 ± 6.3	63.3 ± 5.6
Weight, mean ± SD, kg	69.0 ± 3.0	69.8 ± 2.7	68.0 ± 3.4
Height, mean ± SD, cm	165.0 ± 3.5	165 ± 3.6	166 ± 3.8
BMI, mean ± SD, kg.m^{-2}	25.3 ± 1.4	25.6 ± 1.3	24.9 ± 0.7
IPAQ score			
High, N (%)	3 (12.5)	1 (8.3)	2 (16.6)
Moderate, N (%)	8 (33.3)	5 (41.6)	3 (24.9)
Low, N (%)	13 (54.1)	6 (50)	7 (58.3)

264 IPAQ, International Physical Activity Questionnaire. Differences were not observed between
265 groups (P> .05).

266

267 *3.2. Muscle strength and hypertrophy*

268   Differences were not observed in isometric and dynamic plantar flexor muscle strength
269 between RET and CTR groups (P= .161 − .928; Table 2). Both groups exhibited a similar
270 increase in dynamic muscle strength at 8 weeks (~7 N.m.), but this was only significant in the
271 RET group (6.7 N.m, 95% CI: 0.1 to 13.2, P= .048; Figure 2, panel A), while the RET group
272 presented a significant increase on isometric muscle strength compared to the baseline (14.7
273 N.m, 95% CI: 5.3 to 24.2, P= .006; Figure 2, panel B). Regarding muscle hypertrophy, the RET
274 group exhibited significant improvements in gastrocnemius and soleus muscle thickness, with
275 an adjusted group difference of 0.6 and 0.8mm (P= .006 − .008; Table 2), respectively.
276 Furthermore, an adjusted group difference of 1.3mm in plantar flexors muscle thickness (P=
277 .002; Table 2) was also observed in the RET group with nine participants presenting changes
278 above the minimal difference needed to be considered real (Figure 2, panel C).

279 **Table 2.** Muscle strength, thickness and quality, and functional tests absolute values and change over 8 weeks.

Variables	Baseline	8 weeks	Adjusted mean change			Adjusted group difference			
	Mean ± SD	Mean ± SD	Mean	95% CI	Δ%	Mean	95% CI	P-value	Cohen's d (95% CI)
Plantar flexor muscle strength									
Isometric, N.m									
CTR	98.4 ± 21.9	102.6±29.1	4.2	-6.2 to 14.6	3.8 ± 14.9%	4.6	-2.0 to 11.3	.161	0.6 (-0.2 to 1.4)
RET	108.9 ± 17.6	123.7±24.1	14.7*	5.3 to 24.2	13.6 ± 12.3%				
Dynamic at 30º.sec^{-1}, N.m									
CTR	87.4 ± 28.8	94.2 ± 19.4	6.8	-2.2 to 15.9	13.6 ± 25.6%	-0.2	-5.7 to 5.3	.928	-0.0 (-0.8 to 0.8)
RET	86.6 ± 19.9	93.2 ± 20.3	6.7*	0.1 to 13.2	8.8 ± 11.4%				
Muscle thickness									
Plantar flexors, mm									
CTR	25.5 ± 2.8	25.1 ± 2.9	-0.5	-1.5 to 0.6	-1.7 ± 6.7%	1.3	0.5 to 2.1	**.002**	1.2 (0.5 to 2.3)
RET	25.4 ± 2.4	27.6 ± 3.2	2.1*	0.8 to 3.4	8.4 ± 7.6%				
Gastrocnemius, mm									
CTR	13.0 ± 1.8	12.9 ± 2.5	-0.1	-0.7 to 0.4	-1.7 ± 7.7%	0.6	0.2 to 0.9	**.006**	1.1 (0.2 to 1.9)
RET	12.8 ± 2.2	13.7 ± 2.5	0.9*	0.4 to 1.4	7.3 ± 6.5%				
Soleus, mm									
CTR	12.5 ± 3.3	12.2 ± 3.1	-0.3	-1.1 to 0.4	-1.5 ± 9.9%	0.8	0.2 to 1.3	**.008**	1.0 (0.3 to 2.0)
RET	12.6 ± 2.1	13.8 ± 2.7	1.2*	0.3 to 2.1	9.6 ± 10.2%				
MQ$_{ST}$									
Isometric, N.m.mm^{-1}									
CTR	3.9 ± 0.8	4.1 ± 1.0	0.2	-0.2 to 0.6	5.9 ± 16.1%	0.0	-0.3 to 0.3	.940	0.0 (-0.8 to 0.8)
RET	4.3 ± 0.6	4.5 ± 0.8	0.2	-0.1 to 0.6	5.1 ± 12.2%				
Dynamic, N.m.mm^{-1}									
CTR	3.4 ± 1.1	3.8 ± 0.8	0.3	-0.1 to 0.7	16.4 ± 30.0%	-0.2	-0.4 to 0.1	.129	-0.5 (-1.2 to 0.4)
RET	3.4 ± 0.7	3.4 ± 0.7	0.0	-0.3 to 0.3	0.8 ± 13.4%				
Subcutaneous fat									
Plantar flexors, mm									
CTR	6.8 ± 3.1	6.8 ± 2.7	-0.1	-0.5 to 0.3	1.2 ± 11.9%	-0.1	-0.4 to 0.3	.737	0.0 (-0.8 to 0.8)
RET	6.5 ± 2.7	6.3 ± 2.5	-0.1	-0.8 to 0.5	0.0 ± 14.1%				
MQ$_{EI}$									
Plantar flexors, a.u.									
CTR	20.4 ± 4.7	23.0 ± 5.2	2.6*	0.0 to 5.1	15.1 ± 20.8%	-0.5	-2.0 to 0.9	.468	-0.2 (-1.0 to 0.6)
RET	19.2 ± 3.8	21.0 ± 4.3	1.8	-0.1 to 3.7	10.2 ± 15.3%				
Gastrocnemius, a.u.									

CTR	28.1 ± 7.6	32.6 ± 6.3	4.5*	1.6 to 7.4	20.9 ± 26.1%	-1.7	-3.7 to 0.3	.088	0.6 (-1.4 to 0.2)
RET	27.4 ± 5.9	28.8 ± 6.4	1.3	-1.8 to 4.5	6.0 ± 18.2%				
Soleus, a.u.									
CTR	12.7 ± 5.6	13.3 ± 7.3	0.6	-1.7 to 2.9	6.5 ± 27.2%	0.8	-0.9 to 2.5	.365	0.4 (-0.4 to 1.2)
RET	11.0 ± 4.3	13.2 ± 4.8	2.2	-0.4 to 4.8	28.4 ± 43.6%				
Functional tests									
Stair climb, sec									
CTR	8.7 ± 1.3	8.9 ± 1.1	0.2	-0.3 to 0.6	2.4 ± 7.6%	-0.4	-0.7 to -0.1	**.021**	-1.3 (-2.1 to -0.4)
RET	9.5 ± 2.8	8.7 ± 2.1	-0.8*	-1.2 to -0.3	-6.1 ± 10.1%				
6-m usual walk, sec									
CTR	4.05 ± 0.21	4.11 ± 0.23	0.07	-0.02 to 0.16	1.7 ± 3.5%	-0.05	-0.17 to 0.08	.465	-0.8 (-2.7 to -0.8)
RET	3.97 ± 0.41	3.95 ± 0.55	-0.21	-0.28 to 0.24	-0.5 ± 10.0%				
TUG, sec									
CTR	6.24 ± 0.94	6.41 ± 0.92	0.16	-0.48 to 0.16	3.0 ± 8.2%	-0.1	-0.4 to 0.2	.403	-0.4 (-1.3 to 0.3)
RET	6.58 ± 1.27	6.43 ± 1.17	-0.14	-0.65 to 0.36	-1.5 ± 11.6%				

280   *, Within-groups statistical difference compared to baseline, P<.05. TUG, Timed-up and go test. Cohen's d values of 0.0 to <0.5 indicate small, values of 0.51 to <0.8

281   indicate medium, and values ≥0.8 indicate large effects
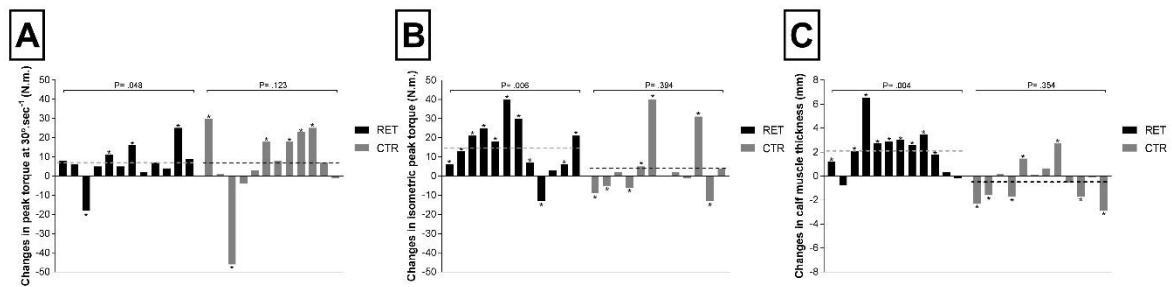
Figure 2. Absolute change in peak torque at $30^{\circ}.sec^{-1}$ (A), isometric peak torque (B) and plantar flexors muscle thickness (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real.

3.3. Muscle quality and functional tests

Both MQ_{EI} and MQ_{ST} measures were not positively affected by the intervention (P= .088 – .940; Table 2). The CTR group experienced a significant increase of 4.5 a.u. on gastrocnemius MQ_{EI} (P= .006), and 2.6 a.u. on plantar flexors MQ_{EI} (P= .048; Figure 3, panel A) at 8 weeks, while no other main effect was detected across time. Likewise, changes in plantar flexors subcutaneous fat thickness were not observed (P= .737; Table 2). Furthermore, changes on MQ_{ST} were not observed across time (P= .086 – .984; Figure 3, panel B). In the functional tests, the RET group presented a significant improvement in stair climb time, with an adjusted group difference of -0.4 sec (P= .021; Figure 3, panel C) with 4 participants presenting changes above the minimal difference needed to be considered real (Figure 3, panel C). No differences in 6-m usual gait time and TUG test performance was observed within- or between-groups (P= .371 – .465; Table 2). No significant association between changes in muscle quality measures and functional capacity tests were found (P= .207 – .815).
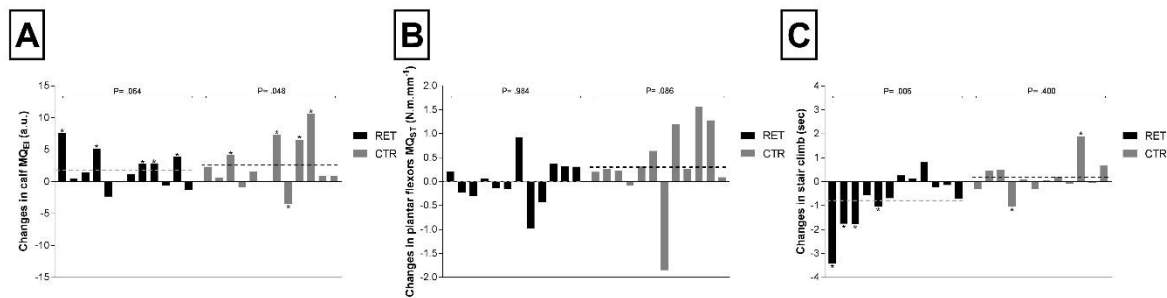
Figure 3. Absolute change in plantar flexors MQ$_{EI}$ (A), dynamic MQ$_{ST}$ (B) and stair climb test (C) for each individual. Black columns, RET group participants; grey columns, CTR group participants; grey dashed lines, average RET group change; black dashed lines, average CTR group change; *, participants presenting changes above the minimal difference needed to be considered real for plantar flexors MQ$_{EI}$ and stair climb test.

4. DISCUSSION

The aims of the present study were to 1) examine the effects of resistance training on muscle quality measures, muscle strength, muscle thickness, and functional capacity, and 2) test if changes in muscle quality and functional capacity performance were associated in older women. We have three important findings. First, both MQ$_{EI}$ and MQ$_{ST}$ do not respond to a short-term intervention, contrary to the findings observed in quadriceps femoris muscles (Wilhelm et al., 2014b; Radaelli et al., 2014; 2019). Secondly, resistance training promotes significant changes in stair climb performance, although it was not in association with changes in plantar flexors-derived muscle quality. Lastly, contrary to our hypothesis, significant gains in plantar flexors muscle hypertrophy were observed after 8 weeks in the RET group. While, although significant changes were observed compared to baseline values, the RET group was not statistically different in plantar flexor muscle strength compared to CTR. Therefore, we expand upon the current knowledge regarding muscle quality and demonstrate that these measures may respond differently in plantar flexor muscles than quadriceps femoris muscles; importantly, muscle quality may not be associated with functional performance, opposite to our expectations.

Regarding muscle composition, although intramuscular lipid stores play a role in providing energy substrates during exercise (Pan et al., 1997), its accumulation in conjunction with increases in fibrous tissue within the muscle are elevated in older adults as a result of

reduced oxidative capacity (Nakagawa et al., 2007). This phenomenon is suggested to impact physical function in older adults (Wilhelm et al., 2014a; Rech et al., 2014; Lopez et al., 2017). It was hypothesized that resistance training may promote changes in plantar flexors derived MQ_{EI} as previously observed in studies examining quadriceps femoris (Wilhelm et al., 2014b; Radaelli et al., 2014; 2019). However, plantar flexors derived MQ_{EI} did not significantly improve following the resistance training program in our study. We would like to suggest some explanations for this. First, it is possible to assume that the baseline values in the present sample were relatively low, despite the lack of studies regarding plantar flexors MQ_{EI} in older adults. Considering that a small effect size could be expected on such low baseline values, although properly calculated, the current sample size may have prevented achieving statistical difference in this outcome. Secondly, the intervention duration itself was shorter than previous studies (8 vs. $12 - 24$ weeks). As we did not know the range or the time-course of changes in this outcome, it is possible to suggest that plantar flexors MQ_{EI} may need longer interventions to identify differences between groups. Thirdly, although quadriceps femoris may respond to a relatively low exercise stimulus as observed in previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014), the proposed number of exercises, intensity or volume (standing calf raise, 3 sets of 6-12 repetitions at 8-15RM) might have been an insufficient stimulus for plantar flexor muscles based on this muscle groups level of activity. Thus, we suggest that plantar flexors MQ_{EI} reductions are more likely to appear following longer periods (at least > 8 weeks) and with greater local stimulus, which may be effective to enhance muscle contraction-induced lipolysis in these muscles (Prats et al., 2006; Fragala et al., 2015).

Curiously, a non-significant association between MQ_{EI} and functional tests' performance was also found in the present study. In contrast to the findings in our present study, Mota et al. (2018) observed a significant negative association between lateral gastrocnemius MQ_{EI} levels and a measure related to functional performance (i.e., plantar flexors rate of velocity development). This would indicate that higher levels of MQ_{EI} in the lateral gastrocnemius may impair the ability to generate velocity rapidly. However, we did not observe similar effects when evaluating functional capacity itself, possibly given the gastrocnemius portion evaluated (medial vs. lateral) or even the study design (RCT vs. cross-sectional study). Therefore, future studies will be necessary to investigate if changes in functional capacity are mediated by such factors in older adults. Altogether, these results suggest that resistance training improves functional capacity regardless of MQ_{EI} adaptations;

365 alternatively, given the lack of changes in MQ_{EI} or measures associated with MQ_{EI}, it was not

366 possible to observe the translation of this outcome to a better functional capacity.

367   While providing significant improvements in the stair climb test, resistance training did

368 not promote meaningful effects on 6-m usual walk and TUG tests, despite a large effect on the

369 6-m usual walk (Cohen's d= 0.8). One of the reasons could be related to the higher baseline

370 values, or even the short intervention duration. In the present study, the participants achieved

371 4.0 and 6.4 sec performance in 6-m usual walk and TUG tests, respectively, while in previous

372 studies the values were 4.4 and 8.4 sec, respectively (Galvão & Taaffe, 2005; Radaelli et al.,

373 2019). The difference between Galvão & Taaffe (2005) and Radaelli et al. (2019) and our study

374 was that the baseline values were superior to the effects that we found following 8 weeks of

375 resistance training (-0.2 and -0.1 sec). Thus, although previous meta-analyses indicate ~0.12

376 m/s increase in gait speed and 0.2 sec decrease in TUG test (Van Abbema et al., 2015), the

377 baseline levels of functional capacity in our sample may have precluded the observation of a

378 larger effect, as previously projected. Furthermore, those previous studies (Galvão & Taaffe,

379 2005; Radaelli et al., 2019) were longer than the present study (i.e., 12 and 20 weeks) which

380 may indicate that at least 12 weeks would be necessary to observe such changes in 6-m usual

381 walk and TUG test. In hindsight, considering that our sample was mostly participants with

382 moderate to lower levels of physical activity who were untrained in resistance exercise,

383 physical adaptations were likely observable but may have required a larger sample size or a

384 longer period of intervention.

385   The significant increase in plantar flexors muscle thickness following a short-term

386 resistance training program was unexpected in the present study. Significant increases of ~2.0

387 mm was found on plantar flexors muscle thickness following 8 weeks of resistance training

388 (n= 9 above the minimal difference needed to be considered real). The reasons for this may be

389 related to the baseline values of our sample. Although the gastrocnemius and soleus muscles

390 are primarily comprised of slow-twitch fibres (i.e., type I) at ~60 and 80% (Gollnick et al.,

391 1974), respectively, the participants presented a gastrocnemius muscle thickness of ~13.0mm,

392 similar to sarcopenic older adults in the studies of Kuyumcu et al. (2016) and Wang et al.

393 (2018) (15.0 and 13.7mm, respectively). Thus, even with an attenuated response in these

394 muscles, given the lower hypertrophic potential compared to fast-twitch fibres (Fry, 2004), the

395 resistance training intervention was likely to induce a significant increase in plantar flexors

396 muscle thickness moderated by the low baseline values. This might explain partially the

397 positive effect on this outcome after a short period of intervention. Furthermore, the significant

increase in muscle thickness was also associated with the lack of changes in MQ_{ST}. Following a short-term resistance training program in untrained participants, we would have expected to observe more neural (i.e., muscle strength) than morphological alterations (i.e., muscle hypertrophy) (Sale, 1988). To the contrary, we observed that the plantar flexor muscles of older women may not respond in that way, resulting in a non-significant change in MQ_{ST} between groups. Thus, dissimilar to results observed in quadriceps femoris muscles (Pinto et al., 2014; Radaelli et al., 2014), it is suggestible that plantar flexor muscles are more likely to present changes in muscle size rather than strength gains following a short-term intervention, particularly when the participants present at baseline with reduced muscle mass levels. Future studies are necessary to elucidate further mechanisms.

Although both groups presented a similar response for MQ_{ST}, MQ_{EI} and muscle strength, only the RET group provided significant improvements on the stair climb functional test. As demonstrated in the present study, the stimuli proposed in the CTR group was enough to promote positive changes in dynamic muscle strength (see Figure 2, panel A). This result was unexpected. However, it could be explained by the frequency of the low-intensity activities in the CTR group and sufficient stimulus on plantar flexor muscles due to the repetitive torque generation provided during treadmill exercise. The same was observed on MIVC where the RET was not superior to the CTR group after 8 weeks. Nevertheless, it is important to note that the CTR protocol did not induce changes in the stair climb test results, which denotes the superiority of resistance training to promote such alterations in physical capacity regardless of changes in plantar flexor muscle strength. This result was interesting as in the study by Suzuki et al. (2001), plantar flexor isometric strength was found to be a significant predictor of the stair climb performance, but not explaining the major part of the variance in that outcome (~2%) (Suzuki et al., 2001). Thus, our results partially agree with Suzuki et al. (2001), demonstrating that although increased following the resistance training program, plantar flexor isometric strength may not account for significant alterations in stair climb test' performance.

The strengths of the present study include expanding the current knowledge about muscle quality and its response to resistance training in a muscle group other than quadriceps femoris in older women. Additionally, the design of an active rather than a pure control group to investigate resistance training effects in this population is novel. However, we have limitations worthy of comment. First, the muscle ultrasound echo intensity was not corrected for intramuscular adipose tissue. Although the ultrasound echo intensity was tested against magnetic resonance imaging in young adults (Young et al., 2015), the equation provided for

medial gastrocnemius was tested in young women (21.9 ± 2.5 yrs) and therefore, inappropriate to be used in older populations given the different age-related factors which affect muscle features (e.g., sarcopenia, decreased fluid storage, and lower glycogen-to-muscle area (Fernández-Elías et al., 2015)). Secondly, the variation in MQ_{EI} values between ultrasound devices makes it difficult to compare different studies. For example, we used a different ultrasound device, setup and data acquisition method to Young et al. (2015). Thus, the design of an imaging phantom in the future may help to adjust MQ_{EI} values of various imaging devices. Thirdly, the sample size required to observe changes in plantar flexors derived outcomes were greater than those previously projected based on quadriceps femoris. Although properly designed, the lack of previous studies investigating the resistance training effects on plantar flexors or plantar flexor muscles hampered a more precise sample size calculation. Future resistance training studies will be able to use the present findings to determine the sample size for these outcomes. Finally, our resistance training prescription did not solely target the plantar flexor muscles and may not have provided a stimulus sufficient to improve MQEI or functional test results. However, previous studies (Radaelli et al., 2014; 2019; Wilhelm et al., 2014b) demonstrated that quadriceps femoris MQEI was likely to present changes in response to single sets of one or two resistance exercises. With this in mind, we targeted multiple muscle groups in accordance with the latest exercise guideline for older adults, focusing on the overall benefits for functional capacity (Fragala et al., 2019) and not exclusively those tested by ourselves.

5. CONCLUSION

In summary, the present study demonstrates that a resistance training program provides significant benefits in the stair climb test, unrelated to plantar flexors-derived muscle quality measures as previously demonstrated in quadriceps femoris (Wilhelm et al., 2014b; Radaelli et al., 2014; Lopez et al., 2017). Furthermore, plantar flexor muscles may also respond differently to a resistance training program, with our cohort of older women presenting significant increases in muscle size rather than muscle strength.

Conflict of Interest: None to declare.

6. REFERENCES

American College of Sports Medicine. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. Med Sci Sports Exerc. 2009 Mar;41(3):687-708.

Arts IM, Pillen S, Schelhaas HJ, Overeem S, Zwarts MJ. Normal values for quantitative muscle ultrasonography in adults. Muscle Nerve. 2010 Jan;41(1):32-41.

Beard JR, Officer AM, Cassels AK. The World Report on Ageing and Health. Gerontologist. 2016 Apr;56 Suppl 2:S163-6.

Buchner DM, de Lateur BJ. The importance of skeletal muscle strength to physical function in older adults. Ann Behav Med. 1991 Sep 1;13(3):91-8.

Capodaglio P, Capodaglio EM, Ferri A, Scaglioni G, Marchi A, Saibene F. Muscle function and functional ability improves more in community-dwelling older women with a mixed-strength training programme. Age Ageing. 2005 Mar;34(2):141-7.

Cohen J. Statistical power analysis. Current directions in psychological science. 1992 Jun;1(3):98-101.

Correa-de-Araujo R, Harris-Love MO, Miljkovic I, Fragala MS, Anthony BW, Manini TM. The Need for Standardized Assessment of Muscle Quality in Skeletal Muscle Function Deficit and Other Aging-Related Muscle Dysfunctions: A Symposium Report. Front Physiol. 2017 Feb 15;8:87.

Fernández-Elías VE, Ortega JF, Nelson RK, Mora-Rodriguez R. Relationship between muscle water and glycogen recovery after prolonged exercise in the heat in humans. Eur J Appl Physiol. 2015;115(9):1919-1926.

Fragala MS, Kenny AM, Kuchel GA. Muscle quality in aging: a multi-dimensional approach to muscle functioning with applications for treatment. Sports Med. 2015 May;45(5):641-58.

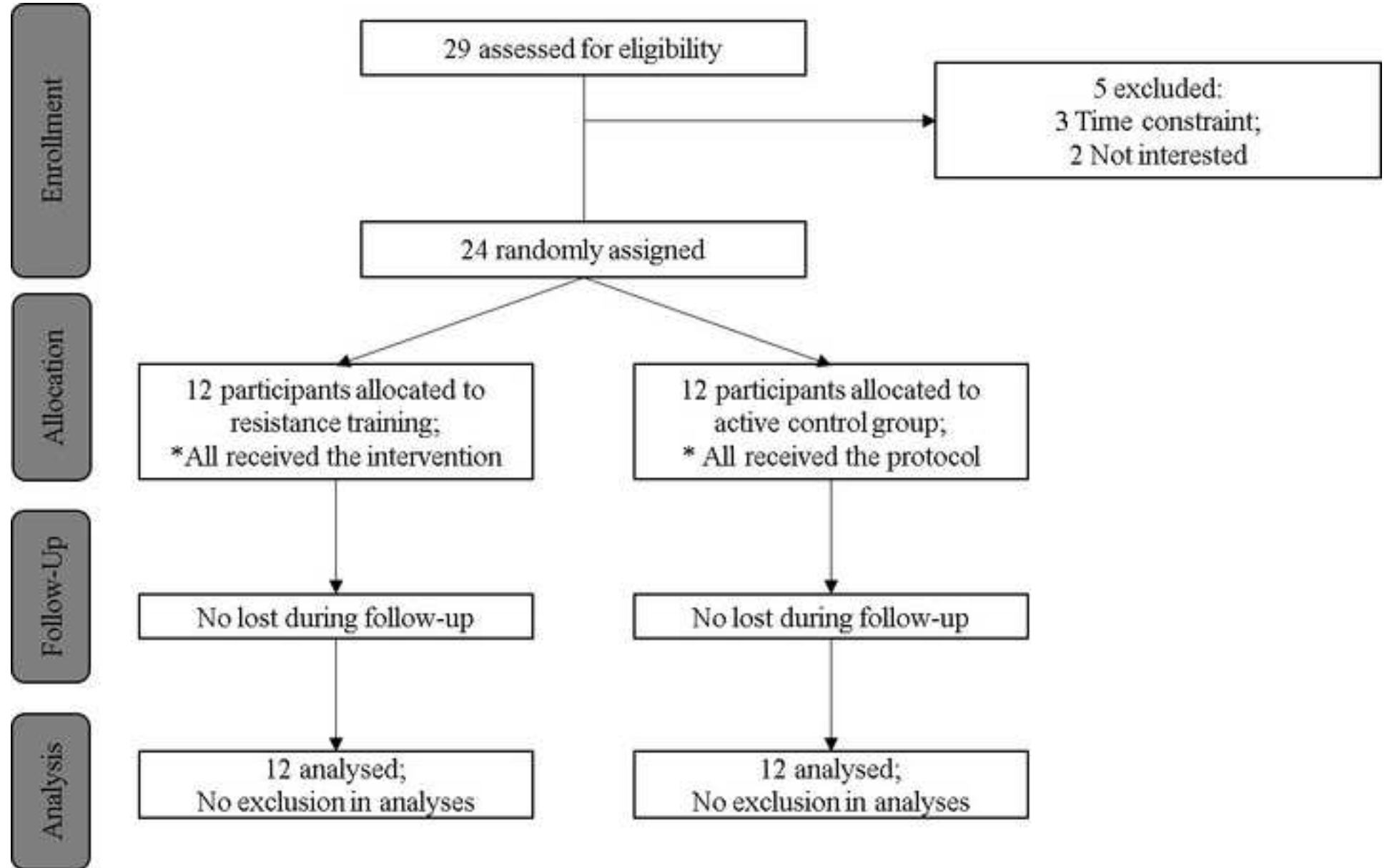Fragala MS, Cadore EL, Dorgo S, Izquierdo M, Kraemer WJ, Peterson MD, Ryan ED. Resistance Training for Older Adults: Position Statement From the National Strength and Conditioning Association. J Strength Cond Res. 2019 Aug;33(8):2019-2052.

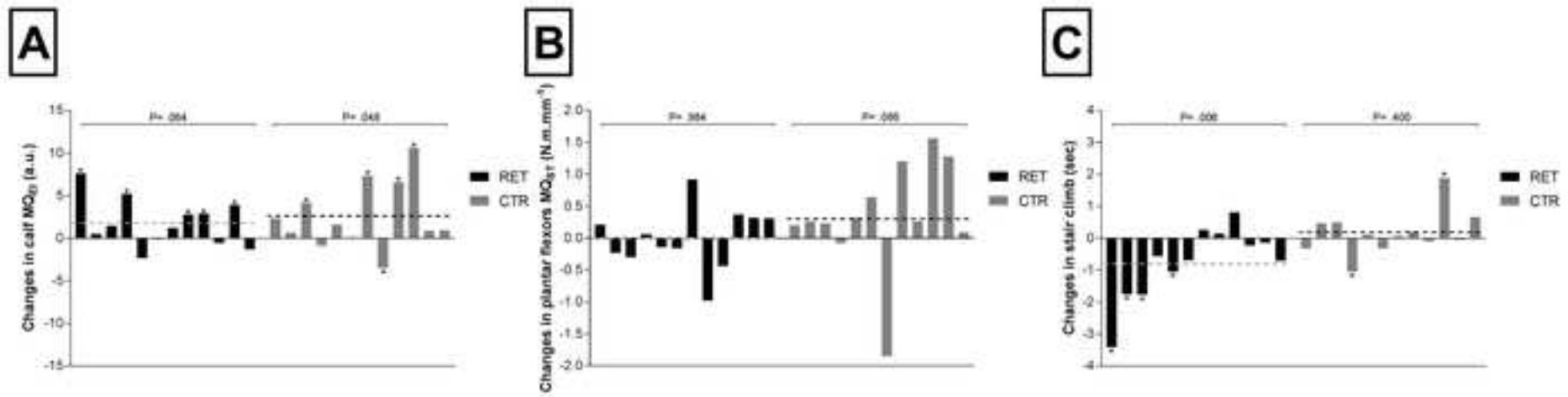Fry AC. The role of resistance exercise intensity on muscle fibre adaptations. Sports Med. 2004;34(10):663-79.

516 Gavin JP, Reeves ND, Jones DA, Roys M, Buckley JG, Baltzopoulos V, Maganaris CN.
517 Combined Resistance and Stretching Exercise Training Benefits Stair Descent Biomechanics
518 in Older Adults. Front Physiol. 2019 Jul 17;10:873.

519 Galvão DA, Taaffe DR. Resistance exercise dosage in older adults: single- versus multiset
520 effects on physical performance and body composition. J Am Geriatr Soc. 2005
521 Dec;53(12):2090-7.

522 Gollnick PD, Sjödin B, Karlsson J, Jansson E, Saltin B. Human soleus muscle: a comparison
523 of fiber composition and enzyme activities with other leg muscles. Pflugers Arch. 1974 Apr
524 22;348(3):247-55.

525 Green J, Forster A, Young J. Reliability of gait speed measured by a timed walking test in
526 patients one year after stroke. Clin Rehabil. 2002 May;16(3):306-14.

527 Kuyumcu ME, Halil M, Kara Ö, Çuni B, Çağlayan G, Güven S, Yeşil Y, Arık G, Yavuz BB,
528 Cankurtaran M, Özçakar L. Ultrasonographic evaluation of the calf muscle mass and
529 architecture in elderly patients with and without sarcopenia. Arch Gerontol Geriatr. 2016 Jul-
530 Aug;65:218-24.

531 Lopez P, Wilhelm EN, Rech A, Minozzo F, Radaelli R, Pinto RS. Echo intensity independently
532 predicts functionality in sedentary older men. Muscle Nerve. 2017 Jan;55(1):9-15.

533 Lopez P, Pinto MD, Pinto RS. Does Rest Time before Ultrasonography Imaging Affect
534 Quadriceps Femoris Muscle Thickness, Cross-Sectional Area and Echo Intensity
535 Measurements? Ultrasound Med Biol. 2019 Feb;45(2):612-616.

536 Lynch NA, Metter EJ, Lindle RS, Fozard JL, Tobin JD, Roy TA, Fleg JL, Hurley BF. Muscle
537 quality. I. Age-associated differences between arm and leg muscle groups. J Appl Physiol
538 (1985). 1999 Jan;86(1):188-94.

539 Mota JA, Giuliani HK, Gerstner GR, Ryan ED. The rate of velocity development associates
540 with muscle echo intensity, but not muscle cross-sectional area in older men. Aging Clin Exp
541 Res. 2018 Jul;30(7):861-865.

542 Nakagawa Y, Hattori M, Harada K, Shirase R, Bando M, Okano G. Age-related changes in
543 intramyocellular lipid in humans by in vivo H-MR spectroscopy. Gerontology.
544 2007;53(4):218-23.

545 Pan DA, Lillioja S, Kriketos AD, Milner MR, Baur LA, Bogardus C, Jenkins AB, Storlien LH.
546 Skeletal muscle triglyceride levels are inversely related to insulin action. Diabetes. 1997
547 Jun;46(6):983-8.

548 Pinto RS, Correa CS, Radaelli R, Cadore EL, Brown LE, Bottaro M. Short-term strength
549 training improves muscle quality and functional capacity of elderly women. Age (Dordr). 2014
550 Feb;36(1):365-72.

551 Prats C, Donsmark M, Qvortrup K, Londos C, Sztalryd C, Holm C, Galbo H, Ploug T. Decrease
552 in intramuscular lipid droplets and translocation of HSL in response to muscle contraction and
553 epinephrine. J Lipid Res. 2006 Nov;47(11):2392-9.

554 Radaelli R, Botton CE, Wilhelm EN, Bottaro M, Lacerda F, Gaya A, Moraes K, Peruzzolo A,
555 Brown LE, Pinto RS. Low- and high-volume strength training induces similar neuromuscular
556 improvements in muscle quality in elderly women. Exp Gerontol. 2013 Aug;48(8):710-6.

557 Radaelli R, Botton CE, Wilhelm EN, Bottaro M, Brown LE, Lacerda F, Gaya A, Moraes K,
558 Peruzzolo A, Pinto RS. Time course of low- and high-volume strength training on
559 neuromuscular adaptations and muscle quality in older women. Age (Dordr). 2014
560 Apr;36(2):881-92.

561 Radaelli R, Brusco CM, Lopez P, Rech A, Machado CLF, Grazioli R, Müller DC, Tufano JJ,
562 Cadore EL, Pinto RS. Muscle quality and functionality in older women improve similarly with
563 muscle power training using one or three sets. Exp Gerontol. 2019 Dec;128:110745.

564 Rech A, Radaelli R, Goltz FR, da Rosa LH, Schneider CD, Pinto RS. Echo intensity is
565 negatively associated with functional capacity in older women. Age (Dordr). 2014;36(5):9708.

566 Sahaly R, Vandewalle H, Driss T, Monod H. Maximal voluntary force and rate of force
567 development in humans--importance of instruction. Eur J Appl Physiol. 2001 Aug;85(3-
568 4):345-50.

569 Sale DG. Neural adaptation to resistance training. Med Sci Sports Exerc. 1988 Oct;20(5
570 Suppl):S135-45.

571 Stephensen D, Drechsler WI, Scott OM. Influence of ankle plantar flexor muscle architecture
572 and strength on gait in boys with haemophilia in comparison to typically developing children.
573 Haemophilia. 2014 May;20(3):413-20.

574 Stock MS, Whitson M, Burton AM, Dawson NT, Sobolewski EJ, Thompson BJ. Echo Intensity
575 Versus Muscle Function Correlations in Older Adults are Influenced by Subcutaneous Fat
576 Thickness. Ultrasound Med Biol. 2018;44(8):1597-1605.

577 Suzuki T, Bean JF, Fielding RA. Muscle power of the ankle flexors predicts functional
578 performance in community-dwelling older women. J Am Geriatr Soc. 2001 Sep;49(9):1161-7.

579 Van Abbema R, De Greef M, Crajé C, Krijnen W, Hobbelen H, Van Der Schans C. What type,
580 or combination of exercise can improve preferred gait speed in older adults? A meta-analysis.
581 BMC Geriatr. 2015 Jul 1;15:72.

582 Wang J, Hu Y, Tian G. Ultrasound measurements of gastrocnemius muscle thickness in older
583 people with sarcopenia. Clin Interv Aging. 2018 Oct 30;13:2193-2199.

584 Wilhelm EN, Rech A, Minozzo F, Radaelli R, Botton CE, Pinto RS. Relationship between
585 quadriceps femoris echo intensity, muscle power, and functional capacity of older men. Age
586 (Dordr). 2014 Jun;36(3):9625.

587 Wilhelm EN, Rech A, Minozzo F, Botton CE, Radaelli R, Teixeira BC, Reischak-Oliveira A,
588 Pinto RS. Concurrent strength and endurance training exercise sequence does not affect
589 neuromuscular adaptations in older men. Exp Gerontol. 2014 Dec;60:207-14.

590 Young HJ, Jenkins NT, Zhao Q, Mccully KK. Measurement of intramuscular fat by muscle
591 echo intensity. Muscle Nerve. 2015 Dec;52(6):963-71.

Figure 1

Click here to access/download;Figure;Figure 1.jpg

Figure 1

Figure 2

Figure 2

Figure 3

Click here to access/download
Table
tables.docx

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Please wait...

If this message is not eventually replaced by the proper contents of the document, your PDF viewer may not be able to display this type of document.

You can upgrade to the latest version of Adobe Reader for Windows®, Mac, or Linux® by visiting http://www.adobe.com/go/reader_download.

For more assistance with Adobe Reader visit http://www.adobe.com/go/acrreader.

Windows is either a registered trademark or a trademark of Microsoft Corporation in the United States and/or other countries. Mac is a trademark of Apple Inc., registered in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.