1. BASIC REPORTING

Regarding the five PeerJ points, i.e., language, context and references, structure, figures, and raw data, the current manuscript meets all. I just have some minor suggestions in their context further presented in general comments of this review.

2. EXPERIMENTAL DESIGN

Among the four PeerJ points about experimental design, the current manuscript presented an original line of investigation and a well-defined research question. In addition, human research ethical was respected. Regarding the methods, further suggestions were provided in general comments of this review.

3. VALIDITY OF FINDINGS

PeerJ listed four points for the validity of findings. The interpretations of findings and some decisions will be critically appraised in general comments of this review. In short, there is some comments and suggestions on results and some speculations, but all aiming to help reach a better level of publication.

4. GENERAL COMMENTS

The manuscript entitled "Declines in skeletal muscle quality vs. size following two weeks of knee joint immobilization" is a very interesting investigation on the magnitude of changes of muscle size and quality after immobilization of lower-limb. The introduction and context were strong enough to picture the dependent variables (i.e., muscle cross-sectional area, and echo intensity) and bring this gap close to the targeted sample. In fact, most part of this knowledge regarding "size and quality" were discussed in the aging field of research, and the authors should be commended thus far to capture this rationale in something really useful in sports medicine such as the ACL injuries. Most parts of my suggestions in the Introduction was minor and just to help improve the real message of this section:

1) In lines 74-77, when the authors stated that ACL injuries were not restricted to athletes, it would be interesting to add a sentence describing what situations the general population may have this injury, e.g. weekend activities or just normal daily activities.

2) In lines 87-88, a reference in the sentence "In recent years, however, investigators have begun to consider changes in muscle quality via analysis of echo intensity" may be necessary.

3) In lines 91-94, an investigation of Nishihara et al. (Clin Interv Aging. 2014 Sep 5;9:1471-8. doi:10.2147/CIA.S67820) provided interesting analyzes about echo intensity frequency and may help to strengthen the sentence.

4) On line 96, when the authors stated that echo intensity is correlated to muscle strength and some functional tests as the chair stand ability, the reference Lopez et al. (Muscle Nerve. 2017 Jan;55(1):9-15. doi: 10.1002/mus.25168) may strengthen the sentence as it found a prediction capacity of echo intensity on 30SS test.

5) In lines 102-104, when the authors have written about "preventative interventions", it could be improved with some evidence that resistance training affects echo intensity in older persons, but its effects in young persons are still lacking, which can suggest your study as a rationale for futures:

Radaelli R, Wilhelm EN, Botton CE, Rech A, Bottaro M, Brown LE, Pinto RS. Effects of single vs. multiple-set short-term strength training in elderly women. Age (Dordr). 2014;36(6):9720. doi: 10.1007/s11357-014-9720-6. Epub 2014 Oct 31.

Radaelli R, Botton CE, Wilhelm EN, Bottaro M, Brown LE, Lacerda F, Gaya A, Moraes K, Peruzzolo A, Pinto RS. Time course of low- and high-volume strength training on neuromuscular adaptations and muscle quality in older women. Age (Dordr). 2014 Apr;36(2):881-92. doi: 10.1007/s11357-013-9611-2.

6) In lines 112-114, in the author's hypothesis, a sentence as "rather than reduction on muscle size" may complete the final phrase.

Regarding the material & methods, I believe that the authors should be also commended due to their transparency and meticulous description of the entire study. Nevertheless, I have a main concern that should be considered as a way to improve more the aforementioned aspect:

7) Please, the authors should inform the sample size calculation, and power reached with the current sample size for the primary outcome (i.e., echo intensity) and secondary.

In the assessment and analyses of the outcomes, I have some concerns mainly in "Ultrasonography Measurements and Analysis". Here, I have some noteworthy points that would like to spend some time:

8) In lines 201-203, the authors informed that the scanning depth was individualized and consistent across trials. As a researcher and reader of muscle ultrasound methods, I understand that some persons present a larger subcutaneous fat that may worsen the acquisition of rectus femoris or vastus intermedius due to the depth capacity of the probe, which is not the case of your study as vastus intermedius was not assessed. At the same time, I also visualize that this choice may implicate in changes in muscle echo intensity due to the specific features and setup of the ultrasound device. I would like to know your opinion about why individualize the muscle scan, and how the authors ensure that the results were not affected by this methodological choice.

9) Before ultrasound images acquisition, the time taken in a supine position is suggested to allow acute re-establishment in body fluids. How long the participants were instructed to keep in supine position before the ultrasound assessment? Some references may help to state this information:

Lopez P, Pinto MD, Pinto RS. Does Rest Time before Ultrasonography Imaging Affect Quadriceps Femoris Muscle Thickness, Cross-Sectional Area and Echo Intensity Measurements? Ultrasound Med Biol. 2019 Feb;45(2):612-616. doi: 10.1016/j.ultrasmedbio.2018.10.010.

Arroyo E, Stout JR, Beyer KS, Church DD, Varanoske AN, Fukuda DH, Hoffman JR. Effects of supine rest duration on ultrasound measures of the vastus lateralis. Clin Physiol Funct Imaging. 2018 Jan;38(1):155-157. doi: 10.1111/cpf.12403.

10) In lines 215-219, I believe that, here, we have some points. The raw echo intensity (based on Fast Fourier Transform – FFT; or "uncorrected" as written in the manuscript) is used in previous literature as a measure of muscle quality. In the study of Young et al. (2015), the values of echo intensity were used to predict intramuscular fat, and the

equation provided was used in the current manuscript, but just for RF. In summary, the authors used each respective method in different muscles, i.e., vastus lateralis echo intensity, and rectus femoris intramuscular fat. Nevertheless, as presented in the main results, behaviors between those variables are also different, i.e., negative and positive correlation between echo intensity and CSA on VL and RF, respectively, and therefore, non-correspondent. Thus, the choice for a corrected and an uncorrected measure makes the main findings sometimes confuse as different things were assessed (echo intensity vs. intramuscular fat; rectus femoris vs. vastus lateralis). Two questions emerge: 1) may the behavior between RF raw echo intensity and intramuscular fat different? 2) the choice for RF intramuscular fat was the reason for its findings? I would like to know the author's opinion and also discuss how to improve the power of synthesis of these results, for example, use for both muscles the raw echo intensity seem to appear like a good way to avoid some misinterpretations.

11) In line 225, just an indication that the percentage values are the coefficient of variations may be necessary for the readers.

Following the methods & materials, actigraphy section was well-described and do not need changes. Regarding the statistical analysis, I would like to weave some comments:

12) The authors choose to test the hypothesis based on the P-values and magnitude-based inferences (MBI). Both methods are often reported in the sports medicine field but sometimes confound what is the reference to ascertain the main findings of the current manuscript as the authors also provided “minimal difference needed to be considered real”. All these 3 parameters may corroborate to certain misinterpretation of findings. What is different? What is higher? What is real? I believe these three questions may drive to organize part of the results section. In order to elaborate on the main results, I suggest focusing on the F and P-values, in addition to the confidence intervals of changes. As we known mean differences and confidence intervals (95% CI) are also indicative of the magnitude of the change. After, maybe on discussion, it would be more interesting to use those different measures to “talk”.

13) In line 255, paired T-test was described as part of the follow-up analyses. In what moment this and Bonferroni test were used? If authors would like to compare baseline values, an independent T-test seems to be a better option to test the difference between legs, and Bonferroni could be used to localize identified differences after ANOVA's as you have 2x2 design.

14) Both η^2 and Cohen's d are measures of effect sizes. Why use two measures of effect size? In most part of the time, both methods indicate the same qualitative intensity. Maybe the choice for one may help to better summarize the results.

In the results section, Table and Figures were adequate to illustrate the results. As the aforementioned suggestions will affect the results, I have just a few comments below:

15) Why posttests comparisons were ? I believe that this approach may not help authors to test the main hypothesis. Also, the use of multiple T-tests may result in higher chances of falls into type I error.

16) After the inspection of the raw data and results section, I believe that are differences between them on RF_{CSA} of the immobilized limb: 6.0±1.17 vs. 6.15±1.17 at baseline (manuscript vs. raw data, respectively). In addition, I do not know what was the method to finding mean differences confidence intervals, or if some adjustment was done (if yes, please, describe in statistical analysis section), but after examining these values through SPSS and excel, both are different from the manuscript (I described the CSA values below, but echo intensity should also be checked). I suggest a double-check in all values presented in manuscript.

RF_{CSA} control	**RF_{CSA} immobilized**
Excel: -0.07 (95% CI, -0.5 to 0.36)	Excel: -0.22 (95% CI, -0.61 to 0.17)
SPSS: -0.07 (95% CI, -0.5 to 0.41)	SPSS: -0.22 (95% CI, -0.65 to 0.21)
Manuscript: 95% CI, -0.46 to 0.63	Manuscript: 95% CI, -0.37 to 0.73

VL_{CSA} control	VL_{CSA} immobilized
Excel: -0.22 (95% CI, -0.99 to 0.56)	Excel: -1.11 (95% CI, -2.16 to -0.05)
SPSS: -0.22 (95% CI, -1.07 to 0.64)	SPSS: -1.11 (95% CI, -2.28 to 0.07)
Manuscript: 95% CI, -0.4 to 0.7	Manuscript: 95% CI, -0.03 to 1.15

17) In lines 266-269, a significant interaction was found in VL echo intensity (F=7.96 and P=0.015). I suggest to double-checking the values of confidence intervals in the immobilized group as its limits are found crossing 0 (95% CI = -5.60 – 0.81) and Bonferroni localized a significant difference in follow-up.

18) Please, it would be better if the authors provided mean changes in some place of the manuscript.

19) As aforementioned, the "minimal difference needed to be considered real – MD" appears in the results section. Unfortunately, is not clear the role of MD. Less than 40% of participants presented some change that exceeds the MD and no mean changes were higher than MD. If the authors pointed a statistical difference in VL echo intensity, and follow, written that the mean change was lower than MD, what is the interpretation? I assume that changes in echo intensity were not real in the cohort and the main conclusion could not indicate a difference in magnitude as it could occur randomly. Please, it would be better for readers to reorganize this section in order to present a clear message.

The discussion of the current manuscript is very interesting. I believe that after our conversations above, many of the aspects between results and discussion will be improved. I just have some few comments in this section:

19) When we see that RF and VL presented different results after immobilization, the first thing that came to mind is: Why? The authors have written that is in accordance with previous studies (Hackney and Ploutz-Snyder 2012 – which is a review, maybe it is better to change studies to comprehensive review), and only in lines 430-431, it is

speculated why it happens. I would like to suggest, maybe, more attention to this point in your discussion.

20) The introduction section is really great and talks about interesting aspects of why investigated this issue in adult women as the incidence of ACL is higher and causes short-to-long periods of immobilization. However, in the discussion, when is expected to see how the results may describe this relationship between immobilization and the context of ACL, it is completed forgot. Maybe, some attention to this point may increase the quality of the discussion and contemplate the proposed rationale.