

Article

Ultrasound Imaging Analysis of the Lumbar Multifidus Muscle Echo Intensity: Intra-Rater and Inter-Rater Reliability of a Novice and an Experienced Rater

Maryse Fortin ^{1,2,3,*}, Brent Rosenstein ¹ , Jerome Levesque ¹ and Neil Nandlall ¹

¹ Department of Health, Kinesiology & Applied Physiology, Concordia University, Montreal, QC H4B 1R6, Canada; brent.rosenstein@concordia.ca (B.R.); levesque.jerome@hotmail.com (J.L.); n_nan@liveconcordia.onmicrosoft.com (N.N.)

² PERFORM Centre, Concordia University, Montreal, QC H4B 1R6, Canada

³ Centre de Recherche Interdisciplinaire en Réadaptation (CRIR), Constance Lethbridge Rehabilitation Centre, Montreal, QC H4B 1T3, Canada

* Correspondence: maryse.fortin@concordia.ca

Abstract: *Background and Objectives:* Ultrasound echo intensity (EI) of the lumbar multifidus muscle (LMM) could offer valuable insights regarding muscle quality in people with low back pain (LBP). However, whether the rater's experience noticeably influences the reliability and precision of LMM EI measurements has not been established. The aims of this study were to investigate the intra-rater and inter-rater reliability of LMM EI measurements, and to compare the reliability and SEM between a novice and an experienced rater. *Materials and Methods:* Twenty athletes (10 females, 10 males) with a history of LBP were included in this study. Transverse ultrasound images of LMM were taken at L5 in prone. LMM EI measurements were obtained bilaterally by tracing the maximum ROI representing the LMM cross-sectional area (CSA), avoiding the inclusion of bone or surrounding fascia. All measurements were performed by two novice raters and an experienced researcher. Each measurement was acquired by each rater three times for each side on three different images, and the average was used in the analyses. Raters were blinded to each other's measurements and the participant's clinical information. Intra-class correlation coefficients (ICCs) were obtained to assess the intra-rater and inter-rater reliability. *Results:* The intra-rater ICC values for the LMM measurements for the experienced rater were excellent (ICC all > 0.997). The inter-rater reliability ICC values showed moderate to excellent reliability (0.614 to 0.994) and agreement between the novice raters and the experienced rater, except for Novice 1 for the right LMM, which revealed lower ICCs and a wider 95% CI. Intra-rater and inter-rater reliability results were similar when separately looking at the right and left side of the muscle and participant gender. *Conclusions:* Our findings support the clinical use of ultrasound imaging for the assessment of LMM EI in individuals with LBP.

Keywords: multifidus muscle; low back pain; ultrasound imaging; echo intensity; reliability



Citation: Fortin, M.; Rosenstein, B.; Levesque, J.; Nandlall, N. Ultrasound Imaging Analysis of the Lumbar Multifidus Muscle Echo Intensity: Intra-Rater and Inter-Rater Reliability of a Novice and an Experienced Rater. *Medicina* **2021**, *57*, 512. <https://doi.org/10.3390/medicina57050512>

Academic Editor: Alberto Tagliacico

Received: 15 March 2021

Accepted: 17 May 2021

Published: 20 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Biomechanical studies have highlighted the important role of the lumbar multifidus muscle (LMM) to provide arthrokinetic control of the vertebral segment, spinal stiffness [1,2], and proprioception of the lumbar spine [3]. Sufficient LMM function is essential to maintain the stability of the kinetic chain and generate forces to the lower and upper limbs [4]. Magnetic resonance imaging (MRI) and ultrasound imaging studies of athletes and non-athletes with LBP have reported morphological changes and functional deficits of the LMM, such as LMM atrophy [5–9], LMM cross-sectional area (CSA) asymmetry [5,7,10–12], increased fatty infiltration [13–16], and increased or decreased muscle activity [17–19]. Magnetic resonance imaging (MRI) remains the gold standard technique for muscle imaging, since the high resolution allows accurate assessment of muscle size and

composition. However, ultrasound imaging is a more accessible and less expensive imaging technique that provides valuable information about muscle function [20–22]. LMM muscle CSA, thickness during submaximal contraction and at rest, and echo intensity (EI) in different postures can be measured with ultrasound to assess muscle size, function, and quality, respectively [23–26].

EI is measured using the ultrasound brightness scale via a gray scale analysis of the pixels in a region of interest (ROI). This measurement can be used as an indicator of muscle quality by estimating intramuscular fat and connective tissue [22,27–29] and muscle damage [30]. Young et al. [22] tested the reproducibility and inter-rater reliability of ultrasound EI in four different muscles and reported both high reproducibility and inter-rater reliability. The authors concluded that ultrasound is an accessible, cost-effective, reproducible imaging technique that is useful to examine skeletal muscle health [22]. This validation study also compared ultrasound EI to MRI, and reported moderate to strong correlations between muscle EI and percent intramuscular fat measured by MRI [22]; this finding is consistent with other studies comparing EI and muscle biopsy samples [31,32]. Similarly, EI of paravertebral muscles in people with chronic low back pain (LBP) is highly correlated with the amount of connective tissue and/or fatty infiltration assessed via biopsy [33]. The increased intramuscular fatty infiltration (e.g., decreased muscle quality) observed in people with chronic LBP [14,15,34] is theorized to increase the risk of injury and reduce overall muscle function [35–37]. Previous studies have also reported that muscle EI is correlated with percentage body fat [23–26], muscle strength and power [38–40], neuromuscular diseases [41,42], and lower cardiovascular performance [38,43,44]. Given these findings, as well as the relative novelty of assessing skeletal muscle EI using ultrasound and its potential clinical use, it is critical to further examine the reliability of this measurement.

Several factors can influence the reliability of EI muscle measurements, including the rater's ultrasound experience, type of muscle, scanning site, and EI measurement method. Although some studies have examined the reliability of EI measurements of the LMM muscle [45–47] with high intra-rater reliability [45,47], none to our knowledge have compared the reliability and standard error of measurements (SEMs) between a novice rater and an experienced rater. Studies suggest that the variability in EI across different muscle groups is a result of the different distribution of fibrous tissue and the orientation of muscle fibers in each muscle group [29,48]. Young et al. [22] reported higher correlations when comparing percent intramuscular fat measured by MRI to corrected muscle EI of each muscle group than when comparing to all muscle groups. Young et al. [22] also found significant variability of muscle EI between two sites of each muscle group. Different EI measurement methods (e.g., size of ROI, imaging plane) influence the reliability of this measure [29,45]. Maximum ROI includes as much muscle as possible, avoiding bone and surrounding fascia [29], while other researchers use rectangular ROI [45]. Sarafraz et al. [45] reported superior intraclass correlation coefficients (ICCs) and SEM values when using maximum ROI in the transverse plane compared to maximum rectangular ROI in the longitudinal plane. Caresio et al. [29] reported moderate to high ICCs for within-session muscle EI depending on ROI size, with larger ROIs leading to higher reliability. Therefore, EI reliability is influenced by ROI size, and there still remains some controversy regarding the reliability of EI muscle measurement in the literature.

An adequate level of reliability is essential to use this ultrasound measure in a clinical setting. More specifically, whether the rater's experience noticeably influences the reliability and precision of LMM ultrasound measurements has not been reported. Given that the LMM plays a key role in lumbopelvic control, the assessment of EI intra-rater and inter-rater reliability warrants further attention. Therefore, the primary aim of this study was to investigate the intra-rater and inter-rater reliability of LMM EI measurements, and to compare the reliability and SEM between a novice and an experienced rater.

2. Materials and Methods

2.1. Participants

Twenty young adults (10 females, 10 males) with a history of LBP from a larger study including soccer, hockey, football, and rugby university level varsity team players were randomly selected and included in the current study. The exclusion criteria included a previous history of severe trauma or spinal fracture, previous spinal surgery, observable spinal abnormalities, and pregnancy, as all of these could affect paraspinal muscle morphology and/or function. This study was approved by the central ethics committee of the Quebec Ministry of Health and Social Services (project #CCER-16-17-06, 21 July 2016). All participants provided informed consent acknowledging that their data would be used for research purposes.

2.2. Procedures

A self-administered questionnaire was completed to collect information regarding participants' demographics and history of LBP. LBP was defined as pain localized between T12 and the gluteal fold [49]. Participants were asked to answer yes or no to the presence of LBP during the past four weeks (pre-season) or three months (off season) prior to the assessment. Players who answered yes to the presence of LBP completed a Numerical Pain Rating Scale (NPRS) to assess average LBP intensity. Participants were also asked about pain location (e.g., center, right side, left side) and pain duration (in months) at both time points. All 20 players in the current study reported LBP in the past four weeks and/or three months prior to measurement.

2.3. Ultrasound

Ultrasound B-mode images of the LMM were captured using a LOGIQ e ultrasound machine (GE Healthcare, Milwaukee, WI, USA) with a 5 MHz curvilinear transducer. The imaging parameters were kept consistent in all acquisitions (frequency: 5 MHz, gain: 60, depth: 8.0 cm).

Participants were placed on a therapy table in a prone position with a pillow under their abdomen to minimize lumbar lordosis [50]. Participants were instructed to relax the paraspinal musculature while the images were obtained. Prior to imaging, the L5 spinous process was palpated and labeled on the skin with a pen. Acoustic coupling gel was administered on the skin, and the ultrasound transducer was positioned longitudinally along the midline of the lumbar spine to detect the location of L5. The transducer was then rotated and positioned transversally over the L5 spinous process for imaging. Transverse images of the LMM at L5 were taken bilaterally, except for larger muscles, where the right and left sides were imaged separately. A total of three images were obtained bilaterally for the right and left LMM.

2.4. Images Assessment

The images were stored and analyzed offline. LMM EI was measured using grayscale analysis imaging via ImageJ software (National Institute of Health, Bethesda, MD, USA, Version 1.49). EI was determined by tracing the maximum ROI representing the LMM cross-sectional area (CSA), avoiding the inclusion of bone or surrounding fascia (Figure 1). EI was then defined as the mean level of gray within the ROI using the grayscale histogram function (e.g., pixels expressed as a value between 0 = black and 255 = white) [27], where enhanced EI is indicative of a greater amount of intramuscular fat and connective tissue. The EI measurements were acquired three times for each side on three different images, and the average was used in the analyses.

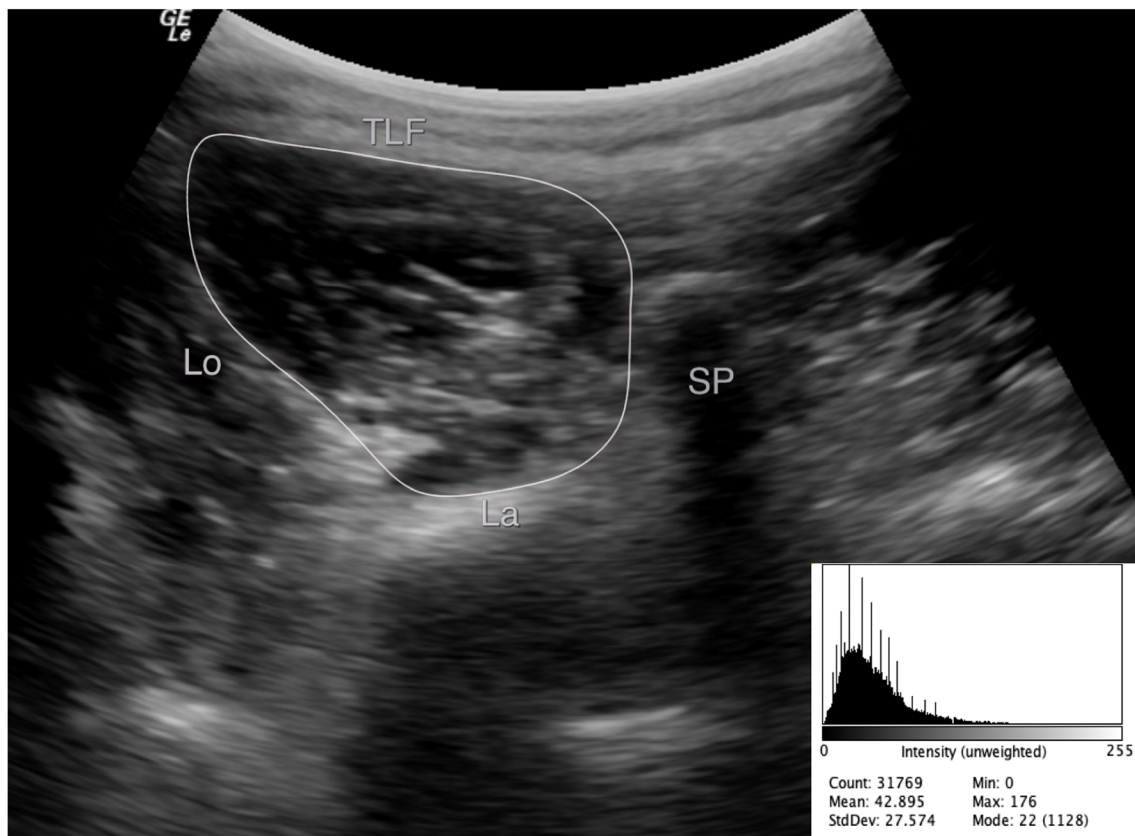


Figure 1. Example of an echo intensity measurement of the left multifidus muscle at L5 from the transverse image and important landmarks used to define the muscle borders. TFL: thoracolumbar fascia, SP: spinous process, La: echogenic laminae, Lo: longissimus muscle.

All muscle measurements were performed by two novice raters (athletic therapy students; Novice 1, N.N. and Novice 2, J.L.) and an experienced researcher (M.F.) with over 10 years of experience in paraspinal imaging analysis. In preparation for this study, the novice raters received training from the experienced rater regarding the anatomy, ultrasound imaging assessment, and how to acquire the LMM CSA and EI measurements. For practice purposes, the novice raters analyzed a sample of five participant images before the start of the measurement study, which were then evaluated and approved by the experienced rater. Then, each rater obtained the muscle measurement three times (on three differing images) on each side, while blinded to each other's measurements and the participant's clinical information. Measurements were obtained once by both novice raters and twice by the experienced rater (with seven days between each time the same image was measured).

2.5. Statistical Analysis

Means and standard deviations were calculated for participants' characteristics and the LMM measurements. The intra-rater reliability for the experienced rater (M.F.), the inter-rater reliability between Novice 1 (N.N.) and the experienced rater (M.F.), and between Novice 2 (J.L.) and the experienced rater (M.F.) were determined by computing the intraclass correlation coefficient (ICC) for the LMM EI measurement. For the intra-rater reliability, the ICC was calculated using a two-way mixed model, average measurement, and absolute agreement. For the inter-rater reliability, the ICC was calculated using a two-way random effects model, average measurement, and absolute agreement. The reliabilities for the right and left LMM were assessed separately. The reliabilities for the male and female participants were assessed as well. The ICCs were interpreted using the following classification, as suggested by Portney and Watkins [51]: less than 0.5 indicates

poor, 0.50–0.74 indicates moderate, 0.75–0.90 indicates good, and greater than 0.90 indicates excellent. The SEM was also calculated to provide an estimate of the expected error related to each muscle measurement. Statistical analysis was performed with the IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Participants

The mean \pm SD age, height, and weight was 21.2 ± 1.3 years, 173.4 ± 10.1 cm, and 75.0 ± 12.3 kg, respectively. A total of 90% ($n = 18$) reported LBP during the pre-season (past four weeks) and 95% ($n = 19$) during the off season (past three months), with a severity of 3.5 ± 1.2 and 4.3 ± 2.0 on the NPRS, respectively. Descriptive data (mean and standard deviation) of participants' baseline characteristics are presented in Table 1.

Table 1. Participants' characteristics.

	All ($n = 20$)	Female ($n = 10$)	Male ($n = 10$)
Age (yr)	21.2 ± 1.3	21.3 ± 1.3	21.0 ± 1.3
Height (cm)	173.4 ± 10.1	166.5 ± 9.1	180.3 ± 5.3
Weight (kg)	75.0 ± 12.3	68.0 ± 11.8	81.9 ± 8.6
BMI (kg/m ²)	24.8 ± 2.6	24.4 ± 2.4	25.3 ± 2.9
Sport, n (%)			
Hockey	11 (55)	5 (50)	6 (60)
Soccer	5 (25)	3 (30)	2 (20)
Rugby	4 (20)	2 (20)	2 (20)
Competitive level (yr) ^a	9.3 ± 4.3	8.6 ± 4.6	10.0 ± 4.0
University level (yr)	2.5 ± 2.2	2.1 ± 1.4	2.8 ± 2.8
LBP four weeks prior, n (%)	18 (90)	9 (90)	9 (90)
LBP three months prior, n (%)	19 (95)	10 (100)	9 (90)
LBP last competitive year, n (%)	12 (60)	6 (60)	6 (60)
LBP location four weeks prior, n (%) ^b			
Centered	5 (28)	2 (22)	3 (33)
Bilateral	6 (33)	3 (33)	3 (33)
Unilateral	7 (39)	4 (44)	3 (33)
LBP location three months prior, n (%) ^a			
Centered	7 (37)	4 (40)	3 (33)
Bilateral	8 (42)	3 (30)	5 (56)
Unilateral	4 (21)	3 (30)	1 (11)
NPRS LBP (0–10) four weeks prior ^b	3.5 ± 1.2	3.5 ± 1.4	3.4 ± 1.1
NPRS LBP (0–10) three months prior ^a	4.3 ± 2.0	3.8 ± 2.3	4.8 ± 1.5

Values are presented as means \pm SDs unless otherwise specified. BMI: body mass index; LBP: lower back pain; NPRS: Numerical Pain Rating Scale. ^a—One missing piece of data from the male group. ^b—One missing piece of data from each male and female group.

3.2. Intra-Rater Reliability

The intra-rater reliability results for LMM measurements of the right and left side for the experienced rater are presented in Table 2. The ICCs indicated excellent intra-rater reliability for all participants, and ranged from 0.997 to 1.000. The SEM values for all participants were relatively small, and ranged between 0.443 and 0.511. The results were similar when separately looking at the right and left side of the muscle, as well as participant gender.

Table 2. Intra-rater reliability of LMM measurements by the experienced rater.

Participants and Side	ICC (95% CI)	SEM
All (<i>n</i> = 20)		
Right	0.999 (0.998–1.000)	0.511
Left	0.999 (0.997–0.999)	0.443
Female (<i>n</i> = 10)		
Right	0.999 (0.996–1.000)	0.466
Left	0.997 (0.986–0.999)	0.592
Male (<i>n</i> = 10)		
Right	0.997 (0.989–0.999)	0.608
Left	0.999 (0.998–1.000)	0.424

ICC: intraclass correlation coefficient; CI: confidence interval; SEM: standard error of measurement.

3.3. Inter-Rater Reliability

The inter-rater reliability results for LMM measurements of the right and left side between the novice raters and the experienced rater are presented in Table 3. The ICCs indicated moderate to excellent inter-rater reliability between Novice 1 and the experienced rater for all participants, and ranged from 0.614 to 0.994. The SEM values for all participants were relatively small, and ranged between 1.999 and 3.308. The results were similar when separately looking at the right and left side of the muscle, as well as participant gender. Notably, the inter-rater reliability between Novice 1 and the experienced rater for the right LMM showed lower ICCs, a wider 95% CI, and larger SEM values. The ICCs indicated excellent inter-rater reliability between Novice 2 and the experienced rater for all participants, and ranged from 0.897 to 0.997. The SEM values for all participants were relatively small, and ranged between 1.432 and 2.257. The results were similar when separately looking at the right and left side of the muscle, as well as participant gender.

Table 3. Inter-rater reliability of LMM measurements between each novice and the experienced rater.

Novice 1 (N.N.) and Experienced Rater			Novice 2 (J.L.) and Experienced Rater		
Participants and Side	ICC (95% CI)	SEM	Participants and Side	ICC (95% CI)	SEM
All (<i>n</i> = 17) ^a			All (<i>n</i> = 20)		
Right	0.963 (0.614–0.991)	3.308	Right	0.992 (0.974–0.997)	1.432
Left	0.982 (0.914–0.994)	1.999	Left	0.972 (0.897–0.990)	2.257
Female (<i>n</i> = 8)			Female (<i>n</i> = 10)		
Right	0.977 (0.372–0.996)	2.351	Right	0.993 (0.963–0.998)	1.237
Left	0.972 (0.661–0.995)	1.854	Left	0.976 (0.895–0.994)	1.707
Male (<i>n</i> = 9)			Male (<i>n</i> = 10)		
Right	0.898 (0.241–0.980)	3.954	Right	0.978 (0.913–0.994)	1.614
Left	0.976 (0.890–0.995)	2.246	Left	0.953 (0.759–0.989)	2.701

^a—One missing piece of data from the male group and two missing data from the female group; ICC: intraclass correlation coefficient; CI: confidence interval; SEM: standard error of measurement.

4. Discussion

The purpose of this study was to investigate the intra-rater and inter-rater reliability of LMM EI measurements using transverse ultrasound images and maximum ROI in 20 participants with a history of LBP, and to compare the reliability and SEM between a novice and an experienced rater. The intra-rater ICC values for the LMM measurements for the experienced rater were all greater than 0.997, indicating excellent reliability. The inter-rater reliability ICC values showed moderate to excellent reliability and agreement between the novice raters and the experienced rater, except for Novice 1 for the right LMM, which revealed lower ICCs and a wider 95% CI. Overall, the intra-rater and inter-rater reliability results were similar when separately looking at the right and left side of the muscle, as well as participant gender.

4.1. Intra-Rater Reliability

Our findings are similar to another study examining the intra-rater reliability of ultrasound LMM EI measurements. Sarafraz et al. [45] investigated within-day intra-rater reliability of several muscles, including the LMM, in 15 participants with a complaint of LBP and 15 healthy controls. The authors measured EI using two orientations, transverse and longitudinal, and captured three ultrasound images in each orientation for each muscle side. In addition, the authors used two different ROI methods: (1) maximum ROI, which includes as much muscle as possible (e.g., the CSA of the LMM in the transverse plane), and (2) a maximum rectangular ROI (longitudinal); the average of the three measurements was used in the analysis. In participants with LBP, using the transverse images and maximum ROI, the authors reported similar ICCs to our study. Their reliability results, however, differed depending on the ROI method and image orientation. Sarafraz et al. [45] reported better reliability with higher ICCs and lower SEM values when using max ROI (e.g., representing the CSA of the LMM) and transverse images in all participants. Similarly, Caresio et al. [29] reported moderate to high ICCs for within-session gastrocnemius and tibialis anterior muscle EI depending on ROI size, with larger ROIs correlating with higher reliability. Therefore, EI reliability is influenced by the ROI size and image orientation, and there still remains some controversy regarding the reliability of LMM EI measurement in the literature.

Yamamoto and Miyazaki [47] investigated between-day (ICCs: right LMM = 0.898 (0.662–0.970), left LMM = 0.775 (0.252–0.934)) and within-day (ICCs: right LMM = 0.981 (0.936–0.994), left LMM = 0.981 (0.935–0.994)) intra-rater reliability of the LMM in 12 healthy participants. Measurements were obtained twice for each side of the LMM on the same day, and on different days using the same maximum ROI definition as our study. Additionally, Resende et al. [46] investigated between-day and within-day intra-rater reliability of the superficial and deep layers of the LMM (e.g., using an ROI box of 1 cm²) in 31 volunteers (e.g., did not report LBP for the past three months) using both transverse and longitudinal images. Interestingly, the authors reported greater ICC values when using bilateral longitudinal images as opposed to transverse images. Overall, our study had higher intra-rater ICCs and lower SEM values than Resende et al. [46]. However, our ROI method included the entire LMM muscle, and made no distinction between the superficial and deep LMM muscle fibers.

Differences between our findings and the previous studies [46,47] could be explained due to methodological differences. In the current study, intra-rater reliability was assessed by repeating the measures on the same ultrasound images seven days apart. Resende et al. [46] and Yamamoto and Miyazaki [47] assessed between-day intra-rater reliability, repeating all procedures on participants during a second visit, which inevitably increased the measurement error. In addition, our study used the average of three measurements from a transverse ultrasound image, while Yamamoto and Miyazaki [47] used the average of two measurements from longitudinal images. As mentioned previously, using different image orientations appears to influence the reliability of EI measurements. Kopenhagen et al. [52] investigated improvements in precision in muscle thickness measurements of the LMM using ultrasound imaging in 30 participants with LBP. The authors reported that, compared to one measurement, the SEM decreased by almost 25% when using an average of two measurements and nearly 50% when using the average of three measurements. The ICC values (95% CI) of the LMM when using a single measurement and the mean of two and three measures also increased from 0.88 (0.76–0.94), 0.94 (0.87–0.97), and 0.96 (0.92–0.98), respectively. The authors stated that measurement precision is optimized by averaging three measures of the LMM. However, averaging more than three measurements led to little or no further improvements in precision [52]. Therefore, it is possible that using an average of three measurements in our study led to a higher reliability and lower SEM compared to averaging two measurements. Furthermore, a study by Resende et al. [46] also investigated the reliability EI measurement of the deep portion of the LMM muscle as opposed to the entire LMM, as measured in the current study. Making

a distinction between the deep and superficial LMM fibers may affect the reliability due to the arbitrary position of the ROI, and the thickness of the LMM will also vary according to individual anthropometric differences. Interestingly, the authors reported higher EI values for the superficial layer than the deep layer. The authors explained that, due to attenuation in the upper parts of the LMM, less energy reaches the deeper layer, making the superficial layer more distinguishable and leading to less variability [46]. Imaging studies have also shown that fatty deposits within the LMM are not homogeneous, and primarily occur in the medial part and deepest layer of the muscle [35,53]. This likely explains why they reported higher ICC values and lower SEM values for the superficial layer compared to the deeper layer.

There are several other factors that can influence the reliability of EI measurements, including the rater's level of experience, type of muscle, and imaging site. Although some studies have examined the reliability of EI measurements of the LMM muscle [45–47] with high intra-rater reliability [45,47], none to our knowledge have compared the reliability and SEMs between a novice rater and an experienced rater. Studies suggest that the variability in EI across different muscle groups is a result of the different distribution of fibrous tissue and the orientation of muscle fibers in each muscle group [29,48]. Young et al. [22] reported higher correlations when comparing percent intramuscular fat measured by MRI to corrected muscle EI of each muscle group than when comparing to all muscle groups. The authors also found significant variability of muscle EI between two sites of each muscle group and percent intramuscular fat between three different MRI slices. This shows the importance of keeping scanning sites consistent, and is in line with the findings of Scholten et al. [54], which emphasized the necessity of measuring the precise muscle location to acquire comparable and reliable EI results across individuals.

An additional factor that can influence the reliability of EI measurements is the participant's gender. To our knowledge, no other studies have separately analyzed the reliability of LMM EI measurements by the participant's gender. However, it is important to note that Fortin et al. [23] and Nandlall et al. [24] showed significantly greater LMM EI values in female young adults than in male young adults when assessing the muscle at rest in a prone position. They concluded that this finding was likely due to females having a greater amount of LMM fatty infiltration/connective tissue due to a naturally higher percentage of body fat, which supports previous findings [14,15,34,55]. A strong correlation was reported between total percent body fat and LMM EI for both male and females [22,23]. Although the current study revealed that LMM EI reliability indices were similar between the male and female participants, this needs to be replicated in the older population.

4.2. Inter-Rater Reliability

Overall, both of the inter-rater ICCs showed moderate to excellent reliability between the novice raters and the experienced rater, except for Novice 1 for the right LMM, which revealed lower ICCs and a wider 95% CI, showing a lack of agreement between the two raters. The results were similar when separately looking at the right and left side of the muscle, as well as participant gender. In addition, our results for both of the novice raters, specifically the right LMM for Novice 1, showed larger SEM values compared to the experienced rater. We are not aware of any other studies that have examined the effect of the rater's level of experience in acquiring LMM EI measurements. However, our findings are in accordance with other studies that have examined EI measures in other muscle groups [56–58].

Ishida et al. [56] investigated the inter-rater reliability of rectus femoris muscle EI measurements in 14 healthy male participants. The authors reported an inter-rater ICC of 0.95 and an SEM of 0.9 between a rater with six years of ultrasound experience and a rater with one year of experience. Rabello et al. [57] investigated the inter-rater reliability of rectus femoris muscle EI measurements in 32 healthy participants (50% female) between a rater with four years of ultrasound experience and a rater with one year of experience. Ultrasound images were obtained at 50% of the rectus femoris belly length and at 70%

of the rectus femoris belly length. At 50% of the rectus femoris belly length, the authors reported inter-rater ICCs of 0.89 (0.82–0.94) and an SEM of 3.95 between the two different experienced raters. At 70% of the rectus femoris belly length, the authors reported inter-rater ICCs of 0.90 (0.82–0.94) and an SEM of 3.87 between the two raters. Therefore, the inter-rater ICCs between the two raters were excellent at each belly length. Lee et al. [58] investigated the inter-rater reliability of abductor pollicis brevis and abductor digiti minimi muscle EI measurements in 20 age-matched controls and 20 patients with carpal tunnel syndrome. The measurements were performed by an experienced rater and a novice rater, but the level of experience for each rater was not mentioned. The authors reported an inter-rater ICC of 0.897 and 0.837 for the abductor pollicis brevis and abductor digiti minimi muscle, respectively. Accordingly, it appears that even with minimal experience or training by the rater, the reliability of EI muscle measurements are good to excellent [56–58]. These findings are similar to our study, which showed good to excellent reliability between the novice raters and the experienced rater, except for Novice 1 for the right LMM, which revealed lower ICCs and a wider 95% CI. Since the novice raters in our study had minimal training by only analyzing a sample of five participants prior to the study, one could expect that more training would have improved the agreement between the two raters.

4.3. Strengths and Limitations

A limitation of this study is the relatively small sample size, which included young and active participants. The reliability of the EI measurements should still be investigated in older and more sedentary individuals, since this factor affects the quality of the ultrasound images. Furthermore, we had no control group. However, our study did investigate individuals with LBP, in which EI was shown to correlate with connective tissue and/or fatty infiltration within muscle tissue in past studies. In addition, to our knowledge, this is the first study to compare the reliability of LMM EI measurements and SEM between a novice and an experienced rater.

5. Conclusions

Our findings suggest that EI LMM measurements using a maximum ROI in the transverse plane have excellent intra-rater reliability (experienced rater) and moderate to excellent reliability between novice and experienced raters. The reliability was also comparable between muscle side (right vs. left) and participant gender. While our results suggest that good reliability was achieved with minimal training from the novice raters, other factors can influence EI reliability measurements, including the type of muscle, imaging site, and ROI size. Our findings support the clinical use of ultrasound imaging for the assessment of LMM EI in individuals with LBP. This measure can offer valuable insights regarding muscle quality of the LMM, which plays a key role in lumbopelvic control.

Author Contributions: Conceptualization, M.F. and B.R.; Methodology, M.F.; Data Collection, M.F., J.L. and N.N.; Formal Analysis, M.F. and B.R.; Data Curation, M.F. and B.R.; Writing—original, B.R. and M.F.; Draft preparation, B.R.; Writing—review and editing, B.R., M.F., J.L. and N.N.; supervision, M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding and was funded internally.

Institutional Review Board Statement: This study was approved by the Central Ethics Committee of the Quebec Ministry of Health and Social Services (project # CCER-16-17-06, 21 July 2016).

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kay, A.G. An extensive literature review of the lumbar multifidus: Biomechanics. *J. Man. Manip. Ther.* **2001**, *9*, 17–39. [[CrossRef](#)]
2. Wilke, H.J.; Wolf, S.; Claes, L.E.; Arand, M.; Wiesend, A. Stability increase of the lumbar spine with different muscle groups. A biomechanical in vitro study. *Spine* **1995**, *20*, 192–198. [[CrossRef](#)] [[PubMed](#)]
3. Brumagne, S.; Cordo, P.; Lysens, R.; Verschuere, S.; Swinnen, S. The role of paraspinal muscle spindles in lumbosacral position sense in individuals with and without low back pain. *Spine* **2000**, *25*, 989–994. [[CrossRef](#)]
4. Kibler, W.B.; Press, J.; Sciascia, A. The role of core stability in athletic function. *Sports Med.* **2006**, *36*, 189–198. [[CrossRef](#)]
5. Barker, K.L.; Shamley, D.R.; Jackson, D. Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: The relationship to pain and disability. *Spine* **2004**, *29*, E515–E519. [[CrossRef](#)] [[PubMed](#)]
6. Danneels, L.A.; Vanderstraeten, G.G.; Cambier, D.C.; Witvrouw, E.E.; De Cuyper, H.J. CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. *Eur. Spine J.* **2000**, *9*, 266–272. [[CrossRef](#)] [[PubMed](#)]
7. Hides, J.; Gilmore, C.; Stanton, W.; Bohlscheid, E. Multifidus size and symmetry among chronic LBP and healthy asymptomatic subjects. *Man. Ther.* **2008**, *13*, 43–49. [[CrossRef](#)]
8. Hides, J.A.; Stanton, W.R.; McMahon, S.; Sims, K.; Richardson, C.A. Effect of stabilization training on multifidus muscle cross-sectional area among young elite cricketers with low back pain. *J. Orthop. Sports Phys. Ther.* **2008**, *38*, 101–108. [[CrossRef](#)]
9. Hides, J.A.; Oostenbroek, T.; Franettovich Smith, M.M.; Mendis, M.D. The effect of low back pain on trunk muscle size/function and hip strength in elite football (soccer) players. *J. Sports Sci.* **2016**, *34*, 2303–2311. [[CrossRef](#)]
10. Kulig, K.; Scheid, A.R.; Beauregard, R.; Popovich, J.M., Jr.; Beneck, G.J.; Colletti, P.M. Multifidus morphology in persons scheduled for single-level lumbar microdiscectomy: Qualitative and quantitative assessment with anatomical correlates. *Am. J. Phys. Med. Rehabil.* **2009**, *88*, 355–361. [[CrossRef](#)]
11. Ploumis, A.; Michailidis, N.; Christodoulou, P.; Kalaitzoglou, I.; Gouvas, G.; Beris, A. Ipsilateral atrophy of paraspinal and psoas muscle in unilateral back pain patients with monosegmental degenerative disc disease. *Br. J. Radiol.* **2011**, *84*, 709–713. [[CrossRef](#)] [[PubMed](#)]
12. Ranger, T.A.; Cicuttini, F.M.; Jensen, T.S.; Peiris, W.L.; Hussain, S.M.; Fairley, J.; Urquhart, D.M. Are the size and composition of the paraspinal muscles associated with low back pain? A systematic review. *Spine J.* **2017**, *17*, 1729–1748. [[CrossRef](#)] [[PubMed](#)]
13. Fortin, M.; Gibbons, L.E.; Videman, T.; Battisti, M.C. Do variations in paraspinal muscle morphology and composition predict low back pain in men? *Scand. J. Med. Sci. Sports* **2015**, *25*, 880–887. [[CrossRef](#)] [[PubMed](#)]
14. Mengiardi, B.; Schmid, M.R.; Boos, N.; Pfirrmann, C.W.; Brunner, F.; Elfering, A.; Hodler, J. Fat content of lumbar paraspinal muscles in patients with chronic low back pain and in asymptomatic volunteers: Quantification with MR spectroscopy. *Radiology* **2006**, *240*, 786–792. [[CrossRef](#)]
15. Sasaki, T.; Yoshimura, N.; Hashizume, H.; Yamada, H.; Oka, H.; Matsudaira, K.; Iwahashi, H.; Shinto, K.; Ishimoto, Y.; Nagata, K.; et al. MRI-defined paraspinal muscle morphology in Japanese population: The wakayama spine study. *PLoS ONE* **2017**, *12*, e0187765. [[CrossRef](#)]
16. Shahidi, B.; Hubbard, J.C.; Gibbons, M.C.; Ruoss, S.; Zlomislic, V.; Allen, R.T.; Garfin, S.R.; Ward, S.R. Lumbar multifidus muscle degenerates in individuals with chronic degenerative lumbar spine pathology. *J. Orthop. Res.* **2017**, *35*, 2700–2706. [[CrossRef](#)] [[PubMed](#)]
17. Larivière, C.; Gagnon, D.; Loisel, P. The comparison of trunk muscles EMG activation between subjects with and without chronic low back pain during flexion-extension and lateral bending tasks. *J. Electromyogr. Kinesiol.* **2000**, *10*, 79–91. [[CrossRef](#)]
18. Van Dieën, J.H.; Cholewicki, J.; Radebold, A. Trunk muscle recruitment patterns in patients with low back pain enhance the stability of the lumbar spine. *Spine* **2003**, *28*, 834–841. [[CrossRef](#)]
19. Van Dieën, J.H.; Selen, L.P.; Cholewicki, J. Trunk muscle activation in low-back pain patients, an analysis of the literature. *J. Electromyogr. Kinesiol.* **2003**, *13*, 333–351. [[CrossRef](#)]
20. Vlychou, M.; Teh, J. Ultrasound of muscle. *Curr. Probl. Diagn. Radiol.* **2008**, *37*, 219–230. [[CrossRef](#)] [[PubMed](#)]
21. Mayans, D.; Cartwright, M.S.; Walker, F.O. Neuromuscular ultrasonography: Quantifying muscle and nerve measurements. *Phys. Med. Rehabil. Clin. N. Am.* **2012**, *23*, 133–148. [[CrossRef](#)] [[PubMed](#)]
22. Young, H.J.; Jenkins, N.T.; Zhao, Q.; McCully, K.K. Measurement of intramuscular fat by muscle echo intensity. *Muscle Nerve* **2015**, *52*, 963–971. [[CrossRef](#)] [[PubMed](#)]
23. Fortin, M.; Rizk, A.; Frenette, S.; Boily, M.; Rivaz, H. Ultrasonography of multifidus muscle morphology and function in ice hockey players with and without low back pain. *Phys. Ther. Sport* **2019**, *37*, 77–85. [[CrossRef](#)] [[PubMed](#)]
24. Nandlall, N.; Rivaz, H.; Rizk, A.; Frenette, S.; Boily, M.; Fortin, M. The effect of low back pain and lower limb injury on lumbar multifidus muscle morphology and function in university soccer players. *BMC Musculoskelet Disord.* **2020**, *21*, 96. [[CrossRef](#)]
25. Schryver, A.; Rivaz, H.; Rizk, A.; Frenette, S.; Boily, M.; Fortin, M. Ultrasonography of lumbar multifidus muscle in university American football players. *Med. Sci. Sports Exerc.* **2020**, *52*, 1495–1501. [[CrossRef](#)]
26. Lévesque, J.; Rivaz, H.; Rizk, A.; Frenette, S.; Boily, M.; Fortin, M. Lumbar multifidus muscle characteristics, body composition, and injury in university rugby players. *J. Athl. Train.* **2020**, *55*, 1116–1123. [[CrossRef](#)]
27. Arts, I.M.; Pillen, S.; Schelhaas, H.J.; Overeem, S.; Zwarts, M.J. Normal values for quantitative muscle ultrasonography in adults. *Muscle Nerve* **2010**, *41*, 32–41. [[CrossRef](#)]

28. Pillen, S.; Tak, R.O.; Zwarts, M.J.; Lammens, M.M.; Verrijp, K.N.; Arts, I.M.; van der Laak, J.A.; Hoogerbrugge, P.M.; van Engelen, B.G.; Verrips, A. Skeletal muscle ultrasound: Correlation between fibrous tissue and echo intensity. *Ultrasound Med. Biol.* **2009**, *35*, 443–446. [[CrossRef](#)]
29. Caresio, C.; Molinari, F.; Emanuel, G.; Minetto, M.A. Muscle echo intensity: Reliability and conditioning factors. *Clin. Physiol. Funct. Imaging* **2015**, *35*, 393–403. [[CrossRef](#)]
30. Chan, R.; Newton, M.; Nosaka, K. Effects of set-repetition configuration in eccentric exercise on muscle damage and the repeated bout effect. *Eur. J. Appl. Physiol.* **2012**, *112*, 2653–2661. [[CrossRef](#)]
31. Reimers, C.D.; Fleckenstein, J.L.; Witt, T.N.; Müller-Felber, W.; Pongratz, D.E. Muscular ultrasound in idiopathic inflammatory myopathies of adults. *J. Neurol. Sci.* **1993**, *116*, 82–92. [[CrossRef](#)]
32. Reimers, K.; Reimers, C.D.; Wagner, S.; Paetzke, I.; Pongratz, D.E. Skeletal muscle sonography: A correlative study of echogenicity and morphology. *J. Ultrasound Med.* **1993**, *12*, 73–77. [[CrossRef](#)] [[PubMed](#)]
33. Eisele, R.; Schmid, R.; Kinzl, L.; Kramer, M.; Katzmaier, P.; Hartwig, E. Soft tissue texture analysis by B-mode-ultrasound in the evaluation of impairment in chronic low back pain. *Eur. J. Ultrasound* **1998**, *8*, 167–175. [[CrossRef](#)]
34. Kjaer, P.; Bendix, T.; Sorensen, J.S.; Korsholm, L.; Leboeuf-Yde, C. Are MRI-defined fat infiltrations in the multifidus muscles associated with low back pain? *BMC Med.* **2007**, *5*, 2. [[CrossRef](#)] [[PubMed](#)]
35. Hildebrandt, M.; Fankhauser, G.; Meichtry, A.; Luomajoki, H. Correlation between lumbar dysfunction and fat infiltration in lumbar multifidus muscles in patients with low back pain. *BMC Musculoskelet Disord.* **2017**, *18*, 12. [[CrossRef](#)] [[PubMed](#)]
36. Le Cara, E.C.; Marcus, R.L.; Dempsey, A.R.; Hoffman, M.D.; Hebert, J.J. Morphology versus function: The relationship between lumbar multifidus intramuscular adipose tissue and muscle function among patients with low back pain. *Arch. Phys. Med. Rehabil.* **2014**, *95*, 1846–1852. [[CrossRef](#)]
37. Masaki, M.; Ikezoe, T.; Fukumoto, Y.; Minami, S.; Tsukagoshi, R.; Sakuma, K.; Ibuki, S.; Yamada, Y.; Kimura, M.; Ichihashi, N. Association of sagittal spinal alignment with thickness and echo intensity of lumbar back muscles in middle-aged and elderly women. *Arch. Gerontol. Geriatr.* **2015**, *61*, 197–201. [[CrossRef](#)]
38. Cadore, E.L.; Izquierdo, M.; Conceição, M.; Radaelli, R.; Pinto, R.S.; Baroni, B.M.; Vaz, M.A.; Alberton, C.L.; Pinto, S.S.; Cunha, G.; et al. Echo intensity is associated with skeletal muscle power and cardiovascular performance in elderly men. *Exp. Gerontol.* **2012**, *47*, 473–478. [[CrossRef](#)]
39. Fukumoto, Y.; Ikezoe, T.; Yamada, Y.; Tsukagoshi, R.; Nakamura, M.; Mori, N.; Kimura, M.; Ichihashi, N. Skeletal muscle quality assessed from echo intensity is associated with muscle strength of middle-aged and elderly persons. *Eur. J. Appl. Physiol.* **2012**, *112*, 1519–1525. [[CrossRef](#)]
40. Watanabe, Y.; Yamada, Y.; Fukumoto, Y.; Ishihara, T.; Yokoyama, K.; Yoshida, T.; Miyake, M.; Yamagata, E.; Kimura, M. Echo intensity obtained from ultrasonography images reflecting muscle strength in elderly men. *Clin. Interv. Aging* **2013**, *8*, 993–998. [[CrossRef](#)]
41. Pillen, S.; Arts, I.M.; Zwarts, M.J. Muscle ultrasound in neuromuscular disorders. *Muscle Nerve* **2008**, *37*, 679–693. [[CrossRef](#)] [[PubMed](#)]
42. Heckmatt, J.Z.; Dubowitz, V.; Leeman, S. Detection of pathological change in dystrophic muscle with B-scan ultrasound imaging. *Lancet* **1980**, *1*, 1389–1390. [[CrossRef](#)]
43. Mangine, G.T.; Fukuda, D.H.; LaMonica, M.B.; Gonzalez, A.M.; Wells, A.J.; Townsend, J.R.; Jajtner, A.R.; Fragala, M.S.; Stout, J.R.; Hoffman, J.R. Influence of gender and muscle architecture asymmetry on jump and sprint performance. *J. Sports Sci. Med.* **2014**, *13*, 904–911. [[PubMed](#)]
44. Mangine, G.T.; Fukuda, D.H.; Townsend, J.R.; Wells, A.J.; Gonzalez, A.M.; Jajtner, A.R.; Bohner, J.D.; LaMonica, M.; Hoffman, J.R.; Fragala, M.S.; et al. Sprinting performance on the Woodway Curve 3.0 is related to muscle architecture. *Eur. J. Sport Sci.* **2015**, *15*, 606–614. [[CrossRef](#)]
45. Sarafraz, H.; Hadian, M.R.; Ayoobi Yazdi, N.; Olyaei, G.; Bagheri, H.; Jalaie, S.; Rasouli, O. Test-retest reliability of nerve and muscle morphometric characteristics utilizing ultrasound imaging in individuals with unilateral sciatica and controls. *Chiropr. Man. Therap.* **2018**, *26*, 47. [[CrossRef](#)] [[PubMed](#)]
46. Resende, E.M.S.; de Oliveira, V.B.; Pereira, W.C.A.; de Oliveira, L.F. Superficial and deep lumbar multifidus layers of asymptomatic individuals: Intraday and interday reliability of the echo intensity measurement. *Eur. Spine J.* **2021**, *30*, 227–231. [[CrossRef](#)]
47. Yamamoto, A.; Miyazaki, J. Examination of the inter-operator reliability of ultrasound measures of multifidus muscle echo intensity. *Rigakuryoho Kagaku* **2018**, *33*, 443–446. [[CrossRef](#)]
48. Pillen, S.; van Alfen, N. Skeletal muscle ultrasound. *Neurol. Res.* **2011**, *33*, 1016–1024. [[CrossRef](#)]
49. Dionne, C.E.; Dunn, K.M.; Croft, P.R.; Nachemson, A.L.; Buchbinder, R.; Walker, B.F.; Wyatt, M.; Cassidy, J.D.; Rossignol, M.; Leboeuf-Yde, C.; et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine* **2008**, *33*, 95–103. [[CrossRef](#)] [[PubMed](#)]
50. Skeie, E.J.; Borge, J.A.; Leboeuf-Yde, C.; Bolton, J.; Wedderkopp, N. Reliability of diagnostic ultrasound in measuring the multifidus muscle. *Chiropr. Man. Therap.* **2015**, *23*, 15. [[CrossRef](#)] [[PubMed](#)]
51. Portney, L.G.; Watkins, M.P. *Foundations of Clinical Research: Applications to Practice*; Prentice Hall: Upper Saddle River, NJ, USA, 2000.

52. Koppenhaver, S.L.; Parent, E.C.; Teyhen, D.S.; Hebert, J.J.; Fritz, J.M. The effect of averaging multiple trials on measurement error during ultrasound imaging of transversus abdominis and lumbar multifidus muscles in individuals with low back pain. *J. Orthop. Sports Phys. Ther.* **2009**, *39*, 604–611. [[CrossRef](#)] [[PubMed](#)]
53. Abbott, R.; Pedler, A.; Sterling, M.; Hides, J.; Murphey, T.; Hoggarth, M.; Elliott, J. The geography of fatty infiltrates within the cervical multifidus and semispinalis cervicis in individuals with chronic whiplash-associated disorders. *J. Orthop. Sports Phys. Ther.* **2015**, *45*, 281–288. [[CrossRef](#)] [[PubMed](#)]
54. Scholten, R.R.; Pillen, S.; Verrips, A.; Zwarts, M.J. Quantitative ultrasonography of skeletal muscles in children: Normal values. *Muscle Nerve* **2003**, *27*, 693–698. [[CrossRef](#)] [[PubMed](#)]
55. Crawford, R.J.; Filli, L.; Elliott, J.M.; Nanz, D.; Fischer, M.A.; Marcon, M.; Ulbrich, E.J. Age- and level-dependence of fatty infiltration in lumbar paravertebral muscles of healthy volunteers. *AJNR Am. J. Neuroradiol.* **2016**, *37*, 742–748. [[CrossRef](#)] [[PubMed](#)]
56. Ishida, H.; Suehiro, T.; Suzuki, K.; Watanabe, S. Muscle thickness and echo intensity measurements of the rectus femoris muscle of healthy subjects: Intra and interrater reliability of transducer tilt during ultrasound. *J. Bodyw. Mov. Ther.* **2018**, *22*, 657–660. [[CrossRef](#)] [[PubMed](#)]
57. Rabello, R.; Fröhlich, M.; Bueno, A.F.; Marcolino, M.A.Z.; De Bona Bernardi, T.; Sbruzzi, G.; Aurélio Vaz, M. Echo intensity reliability between two rectus femoris probe sites. *Ultrasound* **2019**, *27*, 233–240. [[CrossRef](#)]
58. Lee, H.; Jee, S.; Park, S.H.; Ahn, S.C.; Im, J.; Sohn, M.K. Quantitative muscle ultrasonography in carpal tunnel syndrome. *Ann. Rehabil. Med.* **2016**, *40*, 1048–1056. [[CrossRef](#)] [[PubMed](#)]