



HAL
open science

Validity and Reliability of 3-D Ultrasound Imaging to Measure Hamstring Muscle and Tendon Volumes

Antoine Frouin, Hugo Guenanten, Guillaume Le Sant, Lilian Lacourpaille, Martin Liebard, Aurélie Sarcher, Peter J Mcnair, Richard Ellis, Antoine Nordez

► **To cite this version:**

Antoine Frouin, Hugo Guenanten, Guillaume Le Sant, Lilian Lacourpaille, Martin Liebard, et al.. Validity and Reliability of 3-D Ultrasound Imaging to Measure Hamstring Muscle and Tendon Volumes. *Ultrasound in Medicine & Biology*, 2023, 49 (6), pp.1457-1464. 10.1016/j.ultrasmedbio.2023.02.012 . hal-04075502

HAL Id: hal-04075502

<https://hal.science/hal-04075502>

Submitted on 20 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Original Contribution

Validity and Reliability of 3-D Ultrasound Imaging to Measure Hamstring Muscle and Tendon Volumes

Antoine Frouin^{a,b}, Hugo Guenanten^a, Guillaume Le Sant^{a,c}, Lilian Lacourpaille^a,
Martin Liebard^{a,c}, Aurélie Sarcher^a, Peter J. McNair^d, Richard Ellis^{d,e}, Antoine Nordez^{a,d,f,*}

^a Nantes Université, Movement – Interactions – Performance, MIP, UR 4334, F-44000 Nantes, France

^b Institut Sport Atlantique (ISA), Nantes, France

^c School of Physiotherapy, IFM3R, Nantes, France

^d Health and Rehabilitation Research Institute, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

^e Active Living and Rehabilitation: Aotearoa, Health and Rehabilitation Research Institute, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

^f Institut Universitaire de France (IUF), Paris, France

ARTICLE INFO

Keywords:

3-D freehand ultrasound
Hamstring
Volume
Validity
Reliability
Muscle
Tendon
Gel pad
Magnetic resonance imaging

Objective: The validity and reliability of 3-D ultrasound (US) in estimation of muscle and tendon volume was assessed in a very limited number of muscles that can be easily immersed. The objective of the present study was to assess the validity and reliability of muscle volume measurements for all hamstring muscle heads and gracilis (GR), as well as tendon volume for the semitendinosus (ST) and GR using freehand 3-D US.

Methods: Three-dimensional US acquisitions were performed for 13 participants in two distinct sessions on separate days, in addition to one session dedicated to magnetic resonance imaging (MRI). Volumes of ST, semimembranosus (SM), biceps femoris short (BFsh) and long (BFlh) heads, and GR muscles and from the tendon from semitendinosus (STtd) and gracilis (GRtd) were collected.

Results: The bias and the 95% confidence intervals of 3-D US compared with MRI ranged from -1.9 mL (-0.8%) to 1.2 mL (1.0%) for muscle volume and from 0.01 mL (0.2%) to -0.03 mL (-2.6%) for tendon volume. For muscle volume assessed using 3-D US, intraclass correlation coefficients (ICCs) ranged from 0.98 (GR) to 1.00 , and coefficients of variation (CV) from 1.1% (SM) to 3.4% (BFsh). For tendon volume, ICCs were 0.99 , and CVs between 3.2% (STtd) and 3.4% (GRtd).

Conclusion: Three-dimensional US can provide a valid and reliable inter-day measurement of hamstrings and GR for both muscle and tendon volumes. In the future, this technique could be used as an outcome for strengthening interventions and potentially in clinical environments.

Introduction

Appreciation of measurements of muscle and tendon volume is important in understanding human motor function, particularly with respect to adaptations associated with aging [1], training [2,3] and immobilization [4], as well as rehabilitation in clinical settings [5–7]. Concerning the latter, the size of an autogenous hamstring tendon autograft used during the surgical reconstruction of a ruptured anterior cruciate ligament (ACL) is highly predictive of the success/failure of the surgery [8]. In addition, deficits in muscle volume of the semitendinosus (ST) and gracilis (GR) can be observed 2 to 4 y after ACL reconstruction, altering the recovery of force capacity in these muscles [9], potentially playing a role in reduced function. In addition, correlation between ST and GR tendon cross-sectional area (CSA) is highly predictive of their graft diameter [10]. Therefore, quantifying volume of these muscles

could represent a relevant outcome to help clinicians develop rehabilitation/training protocols.

The current gold standard in assessment of muscle and tendon volumes is magnetic resonance imaging (MRI) [10,11]. However, its cost (human/material) and accessibility (time/geography) make it prohibitive in many rehabilitation settings [12]. Freehand 3-D ultrasound (US) imaging combines B-mode US imaging and 3-D motion analysis of the US probe to allow the reconstruction of 3-D images from scans [13]. Three-dimensional US imaging has already been validated against MRI for gastrocnemius medialis muscle. The volume of the distal part of the ST muscle has also been validated and compared with those measured in dissected cadaveric muscles [14]. Changes in ST muscle volume after lengthening surgery in children with spastic paresis have also been reported [14,15]. However, the validity and reliability of 3-D US muscle volume in assessing other hamstring muscles have not been investigated.

* Address correspondence to: Antoine Nordez, Nantes Université, Movement - Interactions - Performance, MIP, UR 4334, F-44000 Nantes, France.

E-mail address: antoine.nordez@univ-nantes.fr (A. Nordez).

Antoine Frouin and Hugo Guenanten contributed equally to this work.

For hamstring tendons, the validity and reliability of their CSA evaluated in 2-D US has been reported [16]. In addition, although the reliability of hamstring tendon volume has recently been reported [17], the validity of such measurements has not been previously reported.

Of note, a main limitation of 3-D US imaging is the physical compression of tissues by the probe leading to distortion in the shape of the underlying tissues [13]. The most commonly used solution to this problem is to scan muscles in the water [13] but this restricts the use of 3-D US imaging to a very limited number of muscles that can be easily immersed (i.e., mainly in the lower leg). A previous study [18] reported that an innovative gel pad could be used to remedy the effects of compression. However, this combination of freehand 3-D US imaging using a gel pad, to quantify muscle and tendon volume, has not been validated against the gold standard (MRI). The applicability of this technique to muscle–tendon groups other than the plantar flexors remains to be demonstrated, and is important to the development of freehand 3-D US imaging for a number of clinical conditions and anatomical areas.

Hence, the primary aim of the current study was to examine the validity of 3-D US imaging in measuring muscle volume of the hamstrings and GR, as well as the tendon volume of the ST and GR using gel pads in healthy participants. A second aim was to examine the test–retest reliability of these measurements. Considering both the hamstring injury rate in sprinting activities [19,20] and the common use of hamstring–gracilis autografts to replace a ruptured ACL [9,21], the present study was focused on the four hamstring heads (i.e., ST, semi-membranosus [SM], biceps femoris short [BFsh] and long heads [BFlh]) and the GR. We hypothesized that 3-D US imaging coupled with a gel pad would be a valid and reliable method to quantify muscle and tendon volumes compared with MRI.

Methods

Participants and experimental design

Thirteen healthy participants (12 men and 1 woman, age 26.6 ± 8.3 y, height 180.3 ± 6.2 cm, mass 73.9 ± 8.3 kg) familiar with weight training volunteered. These participants were recruited from advertisements and word of mouth at a university setting. All participants provided informed consent prior to participation, in accordance with institutional guidelines set by the Declaration of Helsinki (World Declaration of Helsinki, 2013). All procedures were approved by the local ethics committee (CERNI No. 03022021). Potential participants were excluded from the study if they had metal implants, had recently practiced exercises inducing hypertrophy or had any lower limb injury and/or surgery. Participants were asked to avoid physical activities 48 h before and between the imaging sessions. Participants performed three imaging sessions in random order (one MRI and two 3-D US) on separate days within 7 d.

3-D US imaging setup

Two-dimensional B-mode US images were obtained using an US scanner (Aixplorer version 12.3 scanner, SuperSonic Imagine, Aix-en-Provence, France) with a 10–2 linear transducer (40-mm field of view; Vermon, Tours, France) and a 20–6 linear transducer (32-mm field of view, SuperLinear SLH20-6, SuperSonic Imagine, Aix-en-Provence, France) for muscle and for tendon acquisitions, respectively. Image depth was set at 8.5 cm for muscle and 3.5 cm for tendon acquisitions. US images were recorded using a video grabber (ElGato Cam Link, Corsair Components, Fremont, CA, USA). Three-dimensional positions of the transducers were recorded by tracking a 3-D-printed four-marker rigid body attached to the transducers, with an optoelectronic motion capture system collecting at 120 Hz (six cameras, Optitrack Flex 13, NaturalPoint, Corvallis, OR, USA). Data from the US images and from the motion capture system were synchronized and streamed using the open-source software PlusServer (Public Software Library for US Imaging Research, version 2.8.0, Kingston, ON, Canada) [22] and recorded using

the open-source software 3D Slicer (slicer.org, version 4.10.1, Perth, Australia) [22,23].

Temporal calibration, that is, calculation of the latency between the acquisitions of the US images and of their associated positions, was done using the open-source software fCal (Freehand Tracked US Calibration Application, version 2.9.0, Kingston, ON, Canada) [22]. Spatial calibration, that is, calculation of the spatial transformation matrix between the US image plane and the tracker attached to the transducer, was done within 3D Slicer, using a tracked stylus. Both calibrations were done in a water tank for each transducer.

Two gel pads were used to avoid tissue compression and to improve US imaging quality [18]. They were designed and produced to fit appropriately on the posterior thigh depending on the scanning location. The gel pad used for assessing hamstring muscles and GR ($40 \times 60 \times 1.3$ cm) was thicker and larger compared with that used for the tendons ($20 \times 30 \times 0.5$ cm). This was because the curvature was higher and the diameter was smaller in the tendon locations. Experimenters were considerably experienced in the use of the US technique (~ 100 h of familiarization).

3-D US imaging acquisition for muscle volumes

To examine test–retest reliability, participants completed two sessions performed on two different days within a week. Acquisitions involved different limb positions to scan the five aforementioned posterior-medial thigh muscles (Fig. 1). For each position, the limbs of the participants were stabilized comfortably to avoid any movement during the data collection. These positions correspond to the optimal positions for 3-D US scanning found in pilot work. Note that scanning positions slightly differed between acquisitions performed using MRI and US. However, considering that muscles are isovolumic [24], such differences in position are unlikely to bias our measurements. The gel pad was placed over the area to be scanned. Coupling gel was added between the gel pad and the skin and the gel pad surface interface with the US probe to ensure contact between layers.

With respect to positions, while hamstring muscle scans were undertaken in the prone position, participants were placed in a side-lying position for the GR muscle. GR merges proximally with other adductor muscles in a narrow tendon close to the body of the pubis [25]. In addition, scanning the GR proximally can lead to decency/privacy issues, which could have ethical implications. Therefore, we used an anatomical landmark to stop the scan proximally for this muscle. Pilot acquisitions revealed that this method enabled us to maximize the amount of GR volume scanned (70%) for all participants. US imaging was used to identify the anterosuperior iliac spine and the base of the patella. An aluminum-lined tape was placed at 50% of this length and transferred to the medial aspect of the thigh (Fig. 2). In addition, these anatomical landmarks were also easier to locate in MR images for technique comparisons (Fig. 2).

For BFlh and BFsh or ST and SM muscles, multiple sweeps (three to six) were achieved to cover the entire muscle volume, moving the transducer from proximal (the ischial tuberosity for hamstrings and the anatomical landmark for GR) to distal (the pes anserinus for the ST, SM and GR, and the head of the fibula for the BFlh and BFsh) insertions in the transverse plane at a constant speed (~ 1 cm/s). For the GR, two overlapping sweeps were performed to cover the partial muscle volume from the anatomic landmark to the distal insertion.

3-D US imaging acquisition for ST and GR tendon volumes

Participants were placed on their right side (Fig. 1D). The gel pad was placed on the area to be scanned using the same process used during muscle acquisitions. The participants were lying on their side because it avoided gel pad slippage. Furthermore, it also enhanced operator stability in acquiring high-quality images when using the 20/6 probe, which is much more sensitive to tremor.

A – BFlh & BFsh



B – ST & SM



C – GR



D – ST & GR tendons



Figure 1. Participants' positions during freehand 3-D ultrasound scanning process using the gel pad. Positions were stabilized by foam supports so that the participants could remain motionless without having to produce muscle tension. (A) Setup used to scan the biceps femoris long (BFlh) and short (BFsh) head muscles. (B) Setup used to scan the semitendinosus (ST) and semimembranosus (SM) muscles. (C) Setup used to scan the gracilis muscle (GR). (D) The same setup was also used to scan the gracilis and semitendinosus tendons with a thinner gel pad. The orange area corresponds to the sweeping area during the measurement.

The distal tendons of the ST and GR are distally located into the pes anserinus (in association with the sartorius). The depth and echogenicity of both tendons change rapidly, and they are difficult to follow distally. Therefore, the 3-D US acquisition was stopped 6 cm under the musculotendinous junction to stop the acquisitions distally. This junction was

chosen because it represents a noteworthy and reliable anatomical landmark. On the proximal aspect, we wanted to include intramuscular tendon because this volume is physiologically important and could be used, for instance, for ACL grafts. However, the intramuscular tendon becomes thinner and more difficult to visualize and segment proximally (in both



Figure 2. Anatomical landmark used to measure the muscle volume of the gracilis. The anatomical landmark was first determined during the 3-D ultrasound session (A) as 50% of the length between the anterior superior iliac spine and the base of the patella ($L_{ASIS-BP}$). It was transferred to the magnetic resonance imaging scans (B) by identifying (i) the base of the patella, and then by moving up the distance in a straight line to the location of the adhesive strip (ii) where the segmentation could then start proximally.

MRI and 3-D US). Therefore, we also used a landmark to stop tendon acquisitions proximally, and an aluminum-lined tape was placed 6 cm distal to the most distal location of the muscle identified using US. During pilot work, we found that 6 cm represents a good compromise to maintain a significant part of the intramuscular tendon in the analysis and that led to more feasible and accurate segmentation in all our participants' data.

Manual sweeps were performed between proximal and distal landmarks in the transverse plane at a constant speed. One sweep was performed during which the tendon was held in the center of the US screen and two acquisitions were made per tendon.

MRI acquisitions and setup

Participants lay prone on the MRI table with the hip and knee at 0°. Consecutive T1-weighted axial MRIs were recorded of the entire hamstring and GR muscles and tendons from their proximal insertion to their distal anatomical landmark on the lower leg. All participants were scanned using a 3-T magnetic resonance imaging system (Ingenia, Philips, Netherlands) combined with a 3-D e-THRIVE sequence with the following parameters: time to repetition, 6.0 ms; time to echo, 3.0 ms; field of view, 864 × 864 × 472 mm; voxel size, 0.46 × 0.46 × 1 mm; flip angle, 10°. All acquisitions were then imported in 3D Slicer.

3-D US imaging volume reconstruction

For each acquisition, the 3-D volume reconstruction from the 2-D US images was done directly in 3D Slicer, using the module "Volume Reconstruction" [22]. First, a region of interest (ROI) was adapted to the sweeping sequence and was set to be the smallest possible that included the volume of the muscle or tendon in order to minimize the files size while maintaining a high voxel quality. Then, the reconstruction process was automatic and lasted about 3 min for each volume. The algorithm used in 3D Slicer was the same as that used in PlusServer [22]: a 3-D voxel array was filled with the pixels from the US images. The interpolation mode was set to "Linear," meaning that each pixel value was inserted into the spatially nearest set of 8 voxels using trilinear interpolation weights. The compounding mode was set to "Mean", meaning that for each voxel, the resulting pixel value was an average of all inserted pixel values. These parameters allowed for a good resistance to data noise. The size of the voxel was set to 0.10 × 0.10 for the transverse direction and to 1.00 mm for the longitudinal direction, for both muscles and tendons.

Data analysis

Manual segmentation was performed for each muscle and tendon using 3D Slicer, version 4.10.1 [23], for both MRI and US scans. Thus, the same segmentation method was used for both imaging methods and both muscle and tendon acquisitions (Fig. 3).

Muscle boundaries were segmented every 7 mm from proximal to distal insertions, which corresponded to 50–55 slices for the hamstring heads depending on the individual. For the GR, the same proximal anatomic landmark (as that used for US) was identified on MRI scans, and segmentation was initiated at this point (Fig. 2).

Tendon boundaries of the ST and GR were segmented every 3 mm from proximal to distal landmarks, leading to a total of 40 slices. For the reliability examination, the full tendon imaged was segmented between the two landmarks. For the validity examination, a total length of ~12 cm of tendon was segmented (6 cm above and 6 cm below the end of the most distal part of the muscle). Two experienced operators performed the segmentation for muscle and tendon. They were blinded from participant and session.

Statistical analysis

Because data passed the Shapiro–Wilk normality test, parametric procedures were used.

Similarity across 3-D US and MRI was examined using Bland–Altman plots and the 95% limits of agreement analysis [26]. Seven *t*-tests (five muscles, two tendons) were also performed to determine whether the bias was significant or not. For each muscle and tendon volume, test–retest reliability of 3-D US imaging was assessed using the intraclass correlation coefficient $ICC_{(3,1)}$, coefficient of variation (CV) and standard error of measurement (SEM) [27]. The ICC was considered low if ≤ 0.50 , as moderate if in the range 0.50–0.75, as good if ≥ 0.75 and as excellent if ≥ 0.90 [28]. The minimum detectable change (MDC) was calculated as $SEM \times \sqrt{2} \times 1.96$. The statistical significance was set at 0.05. Results are reported as the mean \pm standard deviation.

Results

Validity of 3-D US imaging versus MRI

The muscle volumes measured by 3-D US imaging and MRI were 175.9 ± 37.9 and 176.4 ± 38.7 mL, respectively. The bias between the two methods was -1.9 mL (-0.8%) for SM, -1.3 mL (-0.6%) for BF_{lh}, -0.3 mL (-0.1%) for ST, 0.0 mL (0.0%) for GR and 1.2 mL (1.0%) for BF_{sh} (Fig. 4). The limits of agreements were 4.2 mL (1.8%) for SM, 4.4 mL (2.0%) for ST, 4.7 mL (2.1%) for BF_{lh}, 9.9 mL (7.9%) for BF_{sh} and 7.7 mL (11.9%) for GR. Further evidence of validity was the mean 95% confidence interval (CI), which was 2.4 mL for GR, 2.9 mL for BF_{lh}, 3.1 mL for ST and SM and 3.8 mL for BF_{sh}.

The tendon volumes measured by 3-D US imaging and MRI were 1.73 ± 0.40 and 1.73 ± 0.42 mL, respectively. The bias between the two methods was 0.01 mL (0.2%) for ST_{td} and -0.03 mL (-2.6%) for GR_{td}. The limits of agreement were 0.36 mL (18.3%) and 0.10 mL (9.7%) for ST_{td} and GR_{td}, respectively. Further evidence of validity was the mean 95% CI, which was 0.03 mL and 0.11 mL for ST_{td} and GR_{td}, respectively.

Reliability of 3-D US imaging

Test–retest reliability results are reported in Table 1 [1]. With respect to muscle volume, the ICCs ranged between 0.98 (GR) and 1.00, and the CVs ranged between 1.1% (SM) and 3.4% (BF_{sh}). The SEM and MDC ranged from 1.8 mL (GR) to 3.8 mL (BF_{sh}) and from 5.0 mL (GR) to 10.7 mL (BF_{sh}).

With respect to tendon volume, the ICCs were 0.96 and 0.99, and the CVs ranged between 5.71% and 4.90% for ST and GR, respectively. The SEM and MDC ranged between 0.09 and 0.24 mL for ST_{td} and between 0.04 mL and 0.11 mL for GR_{td}.

Discussion

The present study assessed the reliability and validity of 3-D US imaging coupled with a gel pad to measure the volume of muscles and tendons of the hamstring muscles and GR. The present study revealed that 3-D US, combined with the use of a gel pad, provided hamstring (ST, SM, BF_{lh} and BF_{sh}) and GR muscle–tendon volumes with acceptable bias and limits of agreement when compared with the current gold standard (*i.e.*, MRI). In addition, a good inter-day reliability coefficient was apparent for both muscle and tendon volumes obtained using 3-D US.

Three-dimensional US imaging provides an accurate estimation of hamstring muscles and GR muscle–tendon volumes (*i.e.*, all bias $< 1.0\%$ except that for GR tendon, which was -2.6%) when compared with MRI. The 3-D US method was compared with the current gold standard (MRI), even if this method can also be affected by error of measurements [29]. We acknowledge significant differences in the Bland–Altman plots

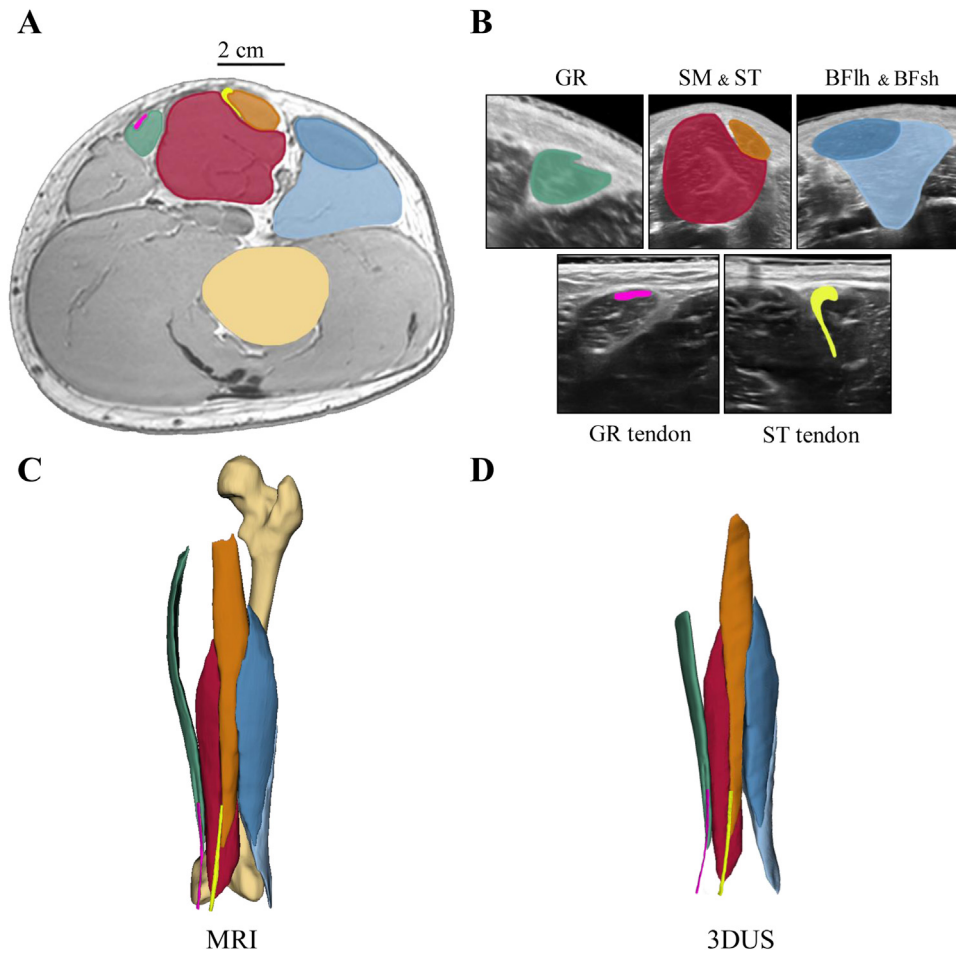


Figure 3. (A) Typical magnetic resonance imaging (MRI) scan with associated segmentation of the five muscles and two tendons. (B) Typical freehand 3-D ultrasound scan (US) of the five acquisitions with associated segmentation of the five muscles and two tendons. (C) Three-dimensional volume reconstruction of the segmented muscles and tendons for MRI (the segmentation of the femur was added for the figure). (D) Three-dimensional volume reconstruction of the segmented muscles and tendons for freehand 3-D US. The biceps femoris long head (BFh) is represented in blue, the biceps femoris short head (BFsh) in light blue, the semitendinosus (ST) in orange, the semimembranosus (SM) in red and the gracilis (GR) in green. Tendons of the GRtd are represented in pink, and those of the STtd in yellow.

for the GR tendon and the BFh muscle. However, the differences were very small, and the bias remained largely acceptable (<3%). Interestingly, limits of agreements were smaller for GR tendon (7.1% and -12.3%) and BFh muscle (4.1% and -5.3%) compared with ST tendon (18.5% and -18.1%) and BFsh (10.9% and -8.9%). Therefore, the fact that the small bias is significant in those two cases does not mean that 3-D US is invalid. Previous studies have reported that freehand 3-D US imaging provides an accurate estimation of gastrocnemius medialis muscle volume [13]. For instance, Barber et al. [13] reported <1.5% differences compared with MRI. Additionally, Haberfehlner et al. [14] found no difference (0.0%) between the distal part of the ST volume assessed by 3-D US and that assessed by dissection. More recently, Devaprakash et al. [30] obtained a bias of 1.1% after comparing volumes obtained by MRI and 3-D US for the Achilles tendon. Accordingly, we found that freehand 3-D US was almost error free compared with MRI in the volume estimation of hamstring muscles. This represents an important step because this muscle group is involved in several clinical scenarios [9], and the work highlights the complex architecture with large intra- and inter-muscular variations [6]. In addition, hamstring muscles are of particular clinical interest in several contexts. For instance, hamstring strain injuries often occur in sports that involve sprinting activity [19]. Monitoring individual hamstring muscle volumes prior to injury could provide valuable information for rehabilitation after an injury. For that instance, it could enable more focused rehabilitation of a specific muscle within the hamstring group [31]. In this context, 3-D US could be very

useful. Our MRI and 3-D US muscle segmentation results are in line with those reported in previous studies for non-sprinters with similar anthropometry parameters (see Additional File S1, Supplementary Material, online only)—65.1 ± 10.9 mL for GR (partial volume), 238.4 ± 45.0 mL for SM, 227.0 ± 55.8 mL for ST, 225.2 ± 53.3 mL for BFh and 125.2 ± 23.6 mL for BFsh—which confirms the consistency of our results.

Recently, Sahinis et al. [17] used a freehand 3-D US imaging technique to estimate hamstring distal tendon volume. For reliability, the authors reported ICCs from 0.96 to 0.98, SEM values from 0.01 to 0.02 mL and MDC values from 0.03 to 0.07 mL (1.5%–3.2%) for ST tendon measurement. Although ICCs were similar, our SEMs and MDCs were slightly higher (Table 1).

In our study, mean ST tendon volume was 1.73 ± 0.41 mL, which was notably higher than reported in previous studies (e.g., 1.44 mL [17] and 1 mL [9]). Two main explanations are proposed for these differences. First, the proximal part of the STtd (above the musculotendinous junction) was not taken into account in these previous studies in which smaller values were observed. Specifically, they segmented the tendon from the distal insertion to the muscle–tendon junction while we segmented up to 6 cm proximally. Our rationale was that this procedure allowed measurement of tendon used in grafts, for example, in ACL reconstruction [32]. Considering that the intramuscular tendon bundle is much more difficult to segment and could lead us to more interpretation, these methodological differences could partly explain the slightly higher SEM and CV obtained in the present study. In addition, we

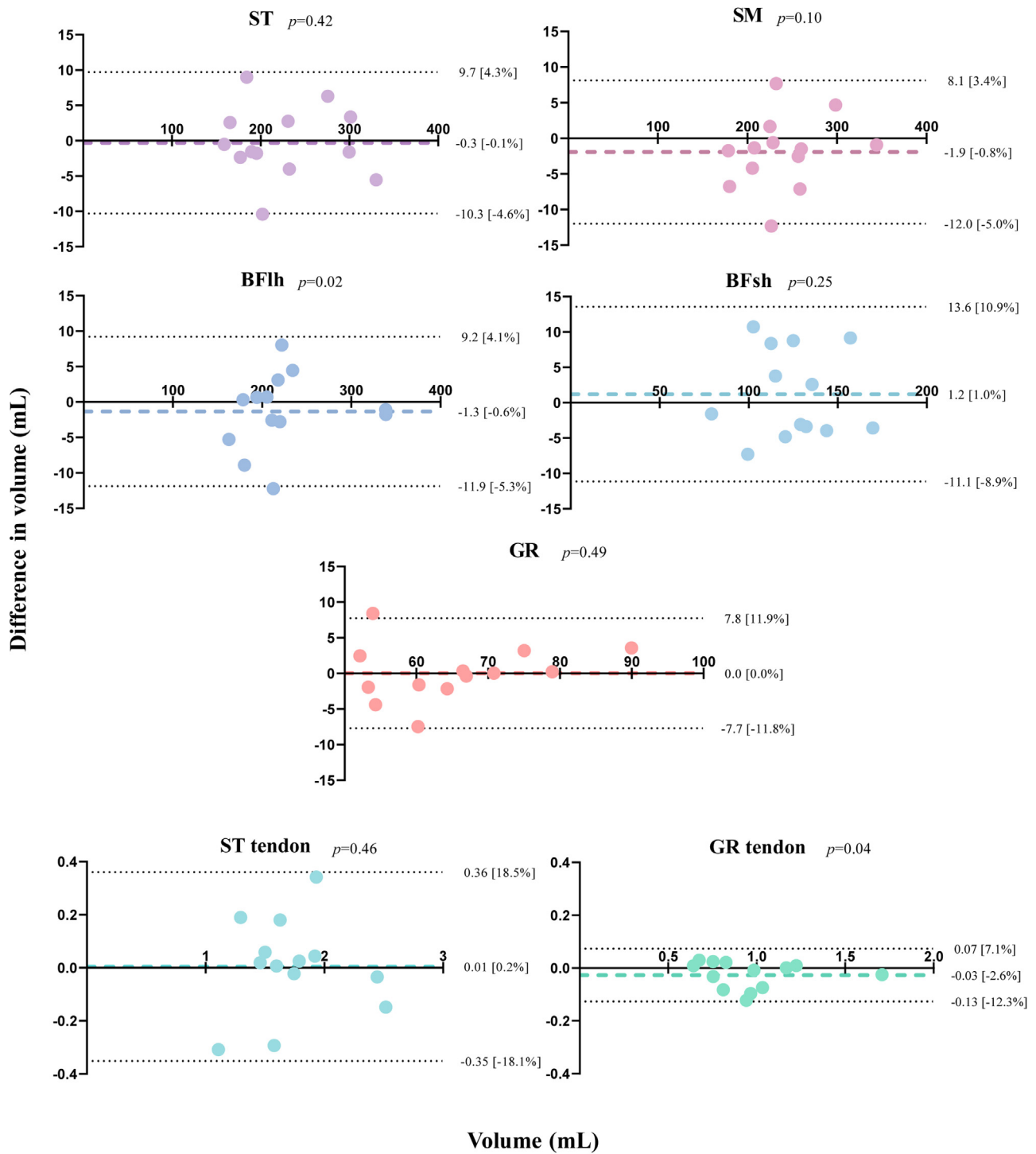


Figure 4. Bland–Altman plots revealing the difference between freehand 3-D ultrasound and magnetic resonance imaging for each muscle and tendon.

realized that this portion is much more likely to have individual variations in terms of morphology (see Additional File S2, Supplementary Material, online only). Further studies might examine the physiological relevance of these variations. In the present study, the tendon volume was not measured up to the insertion distally because we were not able to scan the tendon at its insertion to the pes anserinus. Distal to the medial condyle area the images were too dark to be analyzed with confidence. Second, we reported high inter-individual variability in tendon architecture, in particular in the intramuscular part. This variability could be, at least partially, explained by the training status of our participants. Some participants were involved in strength training programs

and that tendon volume could be positively influenced by weight training [33]. Relatedly, it can also partly explain the differences in volume reported by Sahinis et al. [17]. For instance, MacLeod et al. [34] obtained STtd volumes of 3.68, 2.49 and 2.81 mL for three participants, which are higher than our results, after rehabilitation and 3 y of training in their regular sports.

It should be noted that volumetric measurements require a manual segmentation of the images. The sharpness of the muscular septum was not clearly delineated on each acquisition (see Additional File S3, Supplementary Material, online only). This procedure has been widely acknowledged as time consuming [35]. For instance, the typical

Table 1
Inter-session reliability of muscle and tendon volumes assessed by freehand 3-D ultrasound

Muscle/ tendon	Session 1 (mL)	Session 2 (mL)	ICC	CV (%)	SEM (mL)	MDC (mL)
GR	65.1 ± 11.9	64.7 ± 10.7	0.98	2.77	1.79	4.95
SM	237.7 ± 46.8	237.8 ± 46.3	1.00	1.13	2.75	7.62
ST	226.2 ± 57.4	228.2 ± 57.1	1.00	1.22	2.63	7.28
BFHh	223.7 ± 55.0	225.0 ± 54.2	1.00	1.88	3.81	10.55
BFsh	125.4 ± 24.5	126.1 ± 23.6	0.98	3.37	3.85	10.67
STtd	1.73 ± 0.40	1.73 ± 0.42	0.96	5.71	0.09	0.24
GRtd	0.97 ± 0.29	0.95 ± 0.28	0.99	4.90	0.04	0.11

BFHh, biceps femoris long head; BFsh, biceps femoris short head; CV, coefficient of variation; GR, gracilis; GRtd, gracilis tendon; ICC, intraclass correlation coefficient; MDC, minimum detectable change; SEM, standard error of measurement; SM, semimembranosus; ST, semitendinosus; STtd, semitendinosus tendon.

duration of the segmentation (for both MRI and 3-D US images) for one participant in the present study was 15 h (10 h for the muscles and 5 h for the tendons).

Recent studies have indicated that neural network analyses provide a successful alternative to automatically estimate muscle volume from freehand 3-D US imaging [36] and MRI images [32]. A sufficient database is required to develop and validate this approach for hamstring muscles and tendons. Overall, the use of freehand 3-D US without the need for water and combined with automatic data processing would make this method more feasible in clinical scenarios. Furthermore, as always when using US, experience is very important as US is operator dependent, as already reported for 2-D US and CSA in hamstring muscles [12]. However, in our study, operators had ~5 mo of specific training related to scanning and analyzing the muscles of interest. Therefore, although it is necessary to perform appropriate training, years of experience are not mandatory.

Muscle and tendon volumes are often used to quantify hypertrophy or atrophy over time (e.g., strength training) [2]. For instance, seated leg curl practiced over 12 wk can lead to an improvement of hamstring muscle group volume of 37.8 mL for ST, 16.6 mL for SM, 24.5 mL for BFHh and 7.9 mL for BFsh, on average [37]. As already reported in the quadriceps [38,39], hypertrophy could be inhomogeneous and lead in shape changes of the muscle. Therefore, it remains important to measure several CSAs or to report the muscle volume. A number of studies have estimated the muscle volume based on several CSAs measured using 2-D US and tape measurements. Although this method is easier to implement, it remains less accurate because the motion of the probe is measured with a millimetric precision using motion capture [40,41].

In contrast, a rapid loss of hamstring muscle volume has been observed in bed rest situations [4,42]. Kilroe et al. [4] have also reported a decrease of 156 mL of the whole quadriceps volume after 7 d of disuse caused by immobilization. Furthermore, tendon volume can also be modified by training, with patellar tendon volume increasing after 12 wk of heavy knee extension strength training [43]. If one considers our MDCs concerning muscular volume (range from 3.85 to 10.67 mL), we are confident that the method implemented in the present study provides sufficient accuracy to consider it a good alternative to MRI, which is more expensive to purchase and to service thereafter.

Limitations

Some limitations to the present study that should be considered. First, methodological choices were made for the scanning protocol after the training of experimenters and the pilot experiments. As previously discussed, anatomical landmarks were sometimes used to guide stopping the acquisitions when the quality of scans was not good. An example is during assessment of tendons on the distal and proximal extremities. Additionally, ethical reasons were apparent at times; for example, decency-related modifications had to be implemented when undertaking scans of the GR muscles. In addition, the scanning

position was slightly changed between the various muscle–tendon groups and sometimes slightly differed from the gold standard scans performed using MRI. Considering the isovolumic property of muscles [24], such slight changes should not influence volume measurements. In our experience from pilot testing, the reliability and accuracy were better using these small modifications in positioning. Where possible, we recognize that is very important to prioritize image quality and to perform all the adjustments required for that purpose. Second, although the MRI sequence was optimal or close to optimal for muscles, it was probably not the case for tendons. The size of the voxel in MRI imaging was $0.46 \times 0.46 \times 1.00$ mm. In the current study, the size of the pixel was smaller than that reported in a previous older investigation of ST tendons (0.22×0.22 mm) [44]. Using the high-frequency probe (20–6 MHz) for tendons resulted in a voxel size of $0.10 \times 0.10 \times 1.00$ mm. This could have induced a random error in the comparison between 3-D US and MRI, and the 95% CI would have probably been lower with a better resolution on the MRI imaging. It could also explain partially why the CVs (3.19%–3.43%) were slightly higher in tendon than in muscle.

Conclusion

The present study has found that 3-D US can provide a valid and reliable inter-day measurement of ST, SM, BFHh, BFsh and GR (partial) muscle volume and STtd and GRtd volumes. It is the first study to provide a comprehensive assessment of all muscle within the hamstring muscle group. The error caused by measurement and processing is acceptable and in accordance with previous 3-D US imaging studies. Therefore, it provides evidence for this technique to be used in exercise training and clinical environments (e.g., ACL reconstruction with a hamstring tendon graft or hamstring sprain injury).

Data availability

We performed this study using 3D Slicer software, which is open source. Resources, data, and tutorials for our 3-D ultrasound system are available at <https://github.com/AurelieSar/3Dultrasound/> and <https://www.github.com/AntoineFrouin/3DUS-results> (github.com).

Conflict of interest

The authors declare no competing interests.

Supplementary materials

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.ultrasmedbio.2023.02.012](https://doi.org/10.1016/j.ultrasmedbio.2023.02.012).

References

- [1] Cawthon PM, Travison TG, Manini TM, Patel S, Pencina KM, Fielding RA, et al. Establishing the link between lean mass and grip strength cut points with mobility disability and other health outcomes: proceedings of the Sarcopenia Definition and Outcomes Consortium Conference. *J Gerontol Ser A* 2020;75:1317–23.
- [2] Schoenfeld BJ, Ogborn D, Krieger JW. Dose–response relationship between weekly resistance training volume and increases in muscle mass: a systematic review and meta-analysis. *J Sports Sci* 2017;35:1073–82.
- [3] Lasevicius T, Schoenfeld BJ, Silva-Batista C, Barros T de S, Aihara AY, Brendon H, et al. Muscle failure promotes greater muscle hypertrophy in low-load but not in high-load resistance training. *J Strength Conditioning Res* 2022;36:346–51.
- [4] Kilroe SP, Fulford J, Jackman SR, Van Loon LJC, Wall BT. Temporal muscle-specific disuse atrophy during one week of leg immobilization. *Med Sci Sports Exercise* 2020;52:944–54.
- [5] Dutailis B, Maniar N, Opar DA, Hickey JT, Timmins RG. Lower limb muscle size after anterior cruciate ligament injury: a systematic review and meta-analysis. *Sports Med* 2021;51:1209–26.
- [6] Kellis E. Intra- and inter-muscular variations in hamstring architecture and mechanics and their implications for injury: a narrative review. *Sports Med* 2018;48:2271–83.
- [7] Kellis E, Sahinis C. Is muscle architecture different in athletes with a previous hamstring strain? A systematic review and meta-analysis. *J Funct Morphol Kinesiol* 2022;7:16.
- [8] van Eck CF, Schreiber VM, Mejia HA, Samuelsson K, van Dijk CN, Karlsson J, et al. Anatomic anterior cruciate ligament reconstruction: a systematic review of surgical techniques and reporting of surgical data. *Arthroscopy* 2010;26:S2–12.
- [9] Konrath JM, Vertullo CJ, Kennedy BA, Bush HS, Barrett RS, Lloyd DG. Morphologic characteristics and strength of the hamstring muscles remain altered at 2 years after use of a hamstring tendon graft in anterior cruciate ligament reconstruction. *Am J Sports Med* 2016;44:2589–98.
- [10] Kositsky A, Gonçalves BAM, Stenroth L, Barrett RS, Diamond LE, Saxby DJ. Reliability and validity of ultrasonography for measurement of hamstring muscle and tendon cross-sectional area. *Ultrasound Med Biol* 2020;46:55–63.
- [11] Stenroth L, Sefa S, Arokoski J, Töyräs J. Does magnetic resonance imaging provide superior reliability for Achilles and patellar tendon cross-sectional area measurements compared with ultrasound imaging? *Ultrasound Med Biol* 2019;45:3186–98.
- [12] Franchi MV, Raiteri BJ, Longo S, Sinha S, Narici MV, Csapo R. Muscle architecture assessment: strengths, shortcomings and new frontiers of in vivo imaging techniques. *Ultrasound Med Biol* 2018;44:2492–504.
- [13] Barber L, Barrett R, Lichtwark G. Validation of a freehand 3D ultrasound system for morphological measures of the medial gastrocnemius muscle. *J Biomech* 2009;42:1313–9.
- [14] Haberehner H, Maas H, Harlaar J, Becher JG, Buizer AI, Jaspers RT. Freehand three-dimensional ultrasound to assess semitendinosus muscle morphology. *J Anat* 2016;229:591–9.
- [15] Haberehner H, Jaspers RT, Rutz E, Harlaar J, van der Sluijs JA, Witbreuk MM, et al. Outcome of medial hamstring lengthening in children with spastic paresis: a biomechanical and morphological observational study. *PLoS One* 2018;13:e0192573.
- [16] Galanis N, Savvidis M, Tsifountoudis I, Gkouvas G, Alafropatis I, Kirkos J, et al. Correlation between semitendinosus and gracilis tendon cross-sectional area determined using ultrasound, magnetic resonance imaging and intraoperative tendon measurements. *J Electromyogr and Kinesiol* 2016;26:44–51.
- [17] Sahinis C, Kellis E, Dafkou K, Ellinoudis A. Reliability of distal hamstring tendon length and cross-sectional area using 3-D freehand ultrasound. *Ultrasound Med Biol* 2021;47:2579–88.
- [18] Thomare J, Lacourpaille L, McNair PJ, Crouzier M, Ellis R, Nordez A. A gel pad designed to measure muscle volume using freehand 3-dimensional ultrasonography. *J Ultrasound Med* 2021;40:1245–50.
- [19] Ekstrand J, Waldén M, Hägglund M. Hamstring injuries have increased by 4% annually in men's professional football, since 2001: a 13-year longitudinal analysis of the UEFA Elite Club injury study. *Br J Sports Med* 2016;50:731–7.
- [20] Hickey JT, Timmins RG, Maniar N, Williams MD, Opar DA. Criteria for progressing rehabilitation and determining return-to-play clearance following hamstring strain injury: a systematic review. *Sports Med* 2017;47:1375–87.
- [21] Messer DJ, Shield AJ, Williams MD, Timmins RG, Bourne MN. Hamstring muscle activation and morphology are significantly altered 1–6 years after anterior cruciate ligament reconstruction with semitendinosus graft. *Knee Surg Sports Traumatol Arthrosc* 2020;28:733–41.
- [22] Ungi T, Lasso A, Fichtinger G. Open-source platforms for navigated image-guided interventions. *Med Image Anal* 2016;33:181–6.
- [23] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* 2012;30:1323–41.
- [24] Herbert RD, Bolsterlee B, Gandevia SC. Passive changes in muscle length. *J Appl Physiol* 2019;126:1445–53.
- [25] Khan IA, Bordoni B, Varacallo M. *Anatomy, bony pelvis and lower limb, thigh gracilis muscle*. Treasure Island, FL: StatPearls; 2022.
- [26] Martin Bland J, Altman Douglas G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;327:307–10.
- [27] Hopkins W. *A new view of statistics*. Internet Soc Sport Sci 2000 <http://www.sportsci.org/resource/stats/>.
- [28] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- [29] Pons C, Borotikar B, Garetier M, Burdin V, Ben Salem D, Lempereur M, et al. Quantifying skeletal muscle volume and shape in humans using MRI: a systematic review of validity and reliability. *PLoS One* 2018;13:e0207847.
- [30] Devaprakash D, Lloyd DG, Barrett RS, Obst SJ, Kennedy B, Adams KL, et al. Magnetic resonance imaging and freehand 3-D ultrasound provide similar estimates of free Achilles tendon shape and 3-D geometry. *Ultrasound Med Biol* 2019;45:2898–905.
- [31] Konishi Y, Kinugasa R, Oda T, Tsukazaki S, Fukubayashi T. Relationship between muscle volume and muscle torque of the hamstrings after anterior cruciate ligament lesion. *Knee Surg Sports Traumatol Arthrosc* 2012;20:2270–4.
- [32] Chen X, Lu R, Zhao F. Convolutional neural network-processed MRI images in the diagnosis of plastic bronchitis in children. *Contrast Media Mol Imaging* 2021;2021:1–8.
- [33] Bohm S, Mersmann F, Arampatzis A. Human tendon adaptation in response to mechanical loading: a systematic review and meta-analysis of exercise intervention studies on healthy adults. *Sports Med Open* 2015;1:7.
- [34] Macleod TD, Snyder-Mackler L, Axe MJ, Buchanan TS. Early regeneration determines long-term graft site morphology and function after reconstruction of the anterior cruciate ligament with semitendinosus–gracilis autograft: a case series. *Int J Sports Phys Ther* 2013;8:256–68.
- [35] Le Troter A, Fouré A, Guye M, Confort-Gouny S, Mattei JP, Gondin J, et al. Volume measurements of individual muscles in human quadriceps femoris using atlas-based segmentation approaches. *Magn Reson Mater Phys* 2016;29:245–57.
- [36] Chanti DA, Duque VG, Crouzier M, Nordez A, Lacourpaille L, Mateus D. IFSS-Net: interactive few-shot Siamese network for faster muscle segmentation and propagation in volumetric ultrasound. *IEEE Trans Med Imaging* 2021;40:2615–28.
- [37] Maeo S, Huang M, Wu Y, Sakurai H, Kusagawa Y, Sugiyama T, et al. Greater hamstrings muscle hypertrophy but similar damage protection after training at long versus short muscle lengths. *Med Sci Sports Exercise* 2021;53:825–37.
- [38] Wakahara T, Ema R, Miyamoto N, Kawakami Y. Inter- and intramuscular differences in training-induced hypertrophy of the quadriceps femoris: association with muscle activation during the first training session. *Clin Physiol Funct Imaging* 2017;37:405–12.
- [39] Earp JE, Newton RU, Cormie P, Blazevich AJ. Inhomogeneous quadriceps femoris hypertrophy in response to strength and power training. *Med Sci Sports Exercise* 2015;47:2389–97.
- [40] Mersmann F, Bohm S, Schroll A, Boeth H, Duda G, Arampatzis A. Muscle shape consistency and muscle volume prediction of thigh muscles: thigh muscle volume prediction. *Scand J Med Sci Sports* 2015;25 e208–13.
- [41] Vanmechelen IM, Shortland AP, Noble JJ. Lower limb muscle volume estimation from maximum cross-sectional area and muscle length in cerebral palsy and typically developing individuals. *Clin Biomech* 2018;51:40–4.
- [42] Franchi MV, Sarto F, Simunic B, Pišot R, Narici MV. Early changes of hamstrings morphology and contractile properties during 10 d of complete inactivity. *Med Sci Sports Exercise* 2022;54:1346–54.
- [43] Kongsgaard M, Reitelsheder S, Pedersen TG, Holm L, Aagaard P, Kjaer M, et al. Region specific patellar tendon hypertrophy in humans following resistance training. *Acta Physiol* 2007;191:111–21.
- [44] Hamada M, Shino K, Mitsuoka T, Abe N, Horibe S. Cross-sectional area measurement of the semitendinosus tendon for anterior cruciate ligament reconstruction. *Arthroscopy* 1998;14:696–701.