Journal of Biomechanics 152 (2023) 111567

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech

In vivo 3D muscle architecture quantification based on 3D freehand ultrasound and magnetic resonance imaging

Zhongzheng Wang [a], Antea Destro [a], Sven Petersson [b,c], Francesco Cenni [d], Ruoli Wang [a,*]

[a] KTH MoveAbility Lab, Department of Engineering Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden
[b] Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[c] Department of Medical Radiation Physics and Nuclear Medicine, Karolinska University Hospital, Stockholm, Sweden
[d] Neuromuscular Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

ARTICLE INFO

Keywords:
Tibialis anterior
Gastrocnemius medialis
Fascicle length
Pennation angle
Muscle volume

ABSTRACT

Muscle architecture parameters, such as the fascicle length, pennation angle, and volume, are important muscle morphology characteristics. Accurate *in vivo* quantification of these parameters allows to detect changes due to pathologies, interventions, and rehabilitation trainings, which ultimately impact on muscles' force-producing capacity. In this study, we compared three-dimensional (3D) muscle architecture parameters of the tibialis anterior and gastrocnemius medialis, which were quantified by 3D freehand ultrasound (3DfUS) and a magnetic resonance imaging (MRI) technique, diffusion tensor imaging (DTI), respectively. Sixteen able-bodied subjects were recruited where seven of them received both 3DfUS and MRI measurement, while the rest underwent 3DfUS measurements twice. Good to excellent intra-rater reliability and inter-session repeatability were found in 3DfUS measurements (intra-class correlation coefficient > 0.81). Overall, the two imaging modalities yielded consistent measurements of the fascicle length, pennation angle, and volume with mean differences smaller than 2.9 mm, 1.8°, and 5.7 cm^3, respectively. The only significant difference was found in the pennation angle of the tibialis anterior, although the discrepancy was small. Our study demonstrated, for the first time, that 3DfUS measurement had high reliability and repeatability for measurement of muscle architecture *in vivo* and could be regarded as an alternative to MRI for 3D evaluation of muscle morphology.

1. Introduction

Three-dimensional muscle morphological parameters such as the fascicle length (FL), pennation angle (PA), and muscle volume influence the intrinsic properties of a skeletal muscle, thus impacting the muscle's force-producing capacity (Fukunaga et al., 2001). Muscle architecture might change for various reasons, including pathological conditions. For example, shorter FLs of the biceps and triceps were found in the paretic arm in individuals with post-stroke hemiparesis, which may lead to a limited range of motion in the elbow joint (Nelson et al., 2018). In children with cerebral palsy, muscle volume was significantly correlated with isometric peak torque and joint work (Reid et al., 2015). These morphological parameters are also important parameters for musculoskeletal modelling to estimate individual muscle force and activation level during motion (Charles et al., 2020, Delabastita et al., 2020). More accurate isometric force and ankle torque prediction were achieved in an EMG-driven Hill-type muscle model when using personalized muscle parameters based on ultrasound measurement (De Oliveira and Menegaldo, 2010). Thus, reliable *in vivo* measurement of the skeletal muscle architecture yields essential insights into muscle functions.

Two-dimensional (2D) brightness-mode (B-mode) ultrasound imaging is a widely used technique to measure skeletal muscle architecture *in vivo* due to several advantages such as being radiation-free, good accessibility, low cost, and good temporal resolution. When measuring muscle FL and PA using 2D ultrasound imaging, clear guidelines should be followed to ensure repeatable measurements, for example standardizing measurement location and probe orientation (Klimstra et al., 2007). Within these guidelines, the start and end points of the fascicle should also be clearly visible. Yet, in experimental practice, the fascicles may not be entirely visible due to a limited field-of-view (FOV). Linear trigonometry is often used to estimate the length of the fascicle that is outside the FOV (Brennan et al., 2017). For PA measurement, the

* Corresponding author. KTH MoveAbility Lab, Department of Engineering Mechanics, KTH Royal Institute of Technology, Osquars Backe 18, 10044 Stockholm, Sweden.
E-mail address: ruoli@kth.se (R. Wang).

https://doi.org/10.1016/j.jbiomech.2023.111567
Accepted 23 March 2023
Available online 28 March 2023

ultrasound transducer should be well aligned, with the image plane perpendicular to the deep aponeurosis or in a position where the image quality is optimized (Padhiar et al., 2008). However, misalignment often happens in practice and results in errors in both FL and PA (Bolsterlee et al., 2016). In addition, muscle volume cannot be acquired directly using 2D ultrasound imaging. These factors may limit the application of 2D B-mode ultrasound in muscle architecture quantification.

To overcome the limited FOV of 2D ultrasound, three-dimensional freehand ultrasound (3DfUS) was proposed and has been applied widely in muscle biomechanics research (Barber et al., 2009, Cenni et al., 2018a, Hanssen et al., 2021). 3DfUS combines 2D B-mode ultrasound with a tracking system, such as a 3D optical motion capture system, to identify the positions and orientations of the ultrasound images in a 3D volume (Prager et al., 2002, Cenni et al., 2016). Rana et al. validated the 3D fascicle orientation measurement method using a physical phantom and found that the mean error was less than 0.5° in each plane (Rana and Wakeling, 2011). Skeletal muscle volume estimated based on 3DfUS has been proven to be valid and reliable (Barber et al., 2009). However, no previous study has evaluated the reliability of 3DfUS in muscle FL or PA quantification *in vivo*.

Magnetic resonance imaging (MRI) is another imaging modality that features higher spatial resolution and bigger FOV than ultrasonography. It is considered as the "gold standard" of *in vivo* muscle volume measurement (Mitsiopoulos et al., 1998, Holzbaur et al., 2007). To further quantify muscle FL and PA, one MRI technique, diffusion tensor imaging (DTI) was introduced. In DTI measurement, the diffusion directions of water molecules are tracked, while water molecules tend to diffuse along the long axis of the fascicles (Le Bihan et al., 2001, Oudeman et al., 2016). Therefore, the fascicles can be reconstructed according to the primary diffusion directions in the skeletal muscles. This method was validated in a rat model (Damon et al., 2002) and through microdissection in human skeletal muscle (Bolsterlee et al., 2019). DTI-based measurement has been used to study FL and PA of different human muscles in both healthy and pathological populations (Bolsterlee et al., 2015, D'Souza et al., 2019, Rehmann et al., 2020, Takahashi et al., 2022). In our previous study, the DTI-based method was shown to be a better approach than 2D ultrasound in the quantification of muscle architecture parameters in post-stroke survivors (Korting et al., 2019). However, the DTI-based approach identifies fascicles using tractography and does not directly image the fascicles. Furthermore, the DTI-based method is resource-intensive with limited accessibility compared to ultrasonography. Alternatively, 3DfUS has the potential as an accessibly direct method for the measurement of 3D muscle architecture *in vivo*. Nevertheless, the repeatability and reliability of 3DfUS have not been systematically evaluated.

The objectives of the study were two-fold. The first objective was to evaluate the intra-rater reliability and inter-session repeatability of 3DfUS measurement in quantifying the muscle morphological parameters (FL, PA, and volume) in two major ankle muscles: the tibialis anterior (TA) and gastrocnemius medialis (GM). The second objective was to compare 3DfUS measurement on these muscle parameters to MRI measurement. We hypothesized that there is no significant difference between 3DfUS-based and MRI-based methods in measuring FL, PA, and volume.

2. Materials and methods

2.1. Subjects and experimental design

The study included 16 able-bodied subjects (9 males and 7 females, age 43.8 ± 12.6 years, height 171.6 ± 12.3 cm, weight 73.3 ± 11.5 kg). The subjects were divided into two groups. Group A included 7 subjects (4 males and 3 females) who underwent both MRI and 3DfUS measurements on a randomized side of the legs. Group B consisted of the remaining 9 subjects and underwent only 3DfUS measurements, but twice with a one-week interval on the dominant side. In both 3DfUS and MRI measurements, the subjects were kept in the same joint alignment in a supine position with the thigh relaxing on a wedge and the foot fixed on an MR-compatible foot support (Fig. 1). All subjects participated voluntarily and gave informed written consent before participation. The study was approved by the Swedish Ethical Review Authority (No. 2016/286-32).

2.2. 3DfUS measurements and data processing

B-mode ultrasound images were recorded at 33 Hz using a 38-mm-wide linear transducer (M9, Mindray, Shenzhen, China). A 10-camera optical motion capture system (V16, Vicon, Oxford, UK) was used to track the positions of four reflective markers that were rigidly fixed on the ultrasound transducer. A synchronization device made with piezo crystal was used to trigger the record of the motion capture system while simultaneously sending sound waves to the ultrasound transducer, leading to a visible artifact on the ultrasound image (Weide et al., 2017). Spatial and temporal calibration of the 3DfUS system was performed prior to data acquisition based on a previous study (Cenni et al., 2016). During measurement, a stack of 2D B-mode ultrasound images in the transverse plane of the muscle was acquired by manually moving the ultrasound transducer along the long axis of the target muscle at a stable speed (approximately 1 cm/s). All measurements were performed by one examiner (Z.W.). In each trial, two to three overlapped parallel sweeps were needed to cover the whole muscle. A generous amount of acoustic gel was used to reduce muscle deformation as much as possible. Ultrasound settings such as the gain, depth, and focal depth were optimized for better visualization of muscle borders. An open-source python library (https://gitlab.com/u0078867/py3dfreehandus) was used to process the ultrasound images and motion capture data (Cenni et al., 2016). A 3D image volume was reconstructed by allocating the 2D images in a 3D space with a spatial resolution of 0.2 mm in the transverse plane and 0.5 mm along the sagittal axis (Fig. 2A).

The TA and GM were manually segmented using 3D Slicer (https://www.slicer.org) (Fedorov et al., 2012) and muscle volume was calculated using the surface model generated from the manual segmentation. For FL and PA identification, the sweep covering the central portion of the muscle cross-sectional area was selected from all sweeps of a scan. The plane for fascicle digitization was manually selected as follows. The mid-longitudinal plane of the imaged muscle was first identified and then tilted manually until the aponeurosis and fascicles were clearly visible. The start and end points of 10 fascicles from the distal to the proximal parts of the muscle were manually digitized, and the distances between the two points were defined as FLs. For PA, the angles between the selected fascicle and the deep/superficial aponeurosis were measured. The deep and superficial PAs were averaged and were reported as the PA of the muscle (Fig. 2B). As the TA has a bipennated architecture, the same procedure was repeated for each subcompartment of the muscle (Fig. 2B).

2.3. MRI protocol and fascicle reconstruction

All subjects in Group A were scanned using a 3.0-Tesla MRI scanner (Prisma Fit, Siemens, Erlangen, Germany). T1-weighted images were acquired using a Turbo-Spin-Echo sequence with the following settings: TR/TE 605/23 ms, FOV 430 mm, acquisition matrix 512 × 299, voxel size 0.84 × 0.84 × 5 mm^3, and scan time 101 s. The DTI images were acquired using an Echo-Planar-Imaging sequence with the following settings: TR/TE 6100/63 ms, FOV 350 mm, acquisition matrix 140 × 140, voxel size 2.5 × 2.5 × 2.5 mm^3, 20 DTI gradient directions, number of signal averages 4, b = 500 s/mm^2 (B0 image with b = 0 s/mm^2, with 2 signal averages), and scan time 520 s.

The processes of fascicle reconstructions were modified based on our previous study (Korting et al., 2019). Briefly, the muscle was first manually segmented based on the T1-weighted images and a 3D mesh surface model was generated (Fig. 3A). The muscle volume was

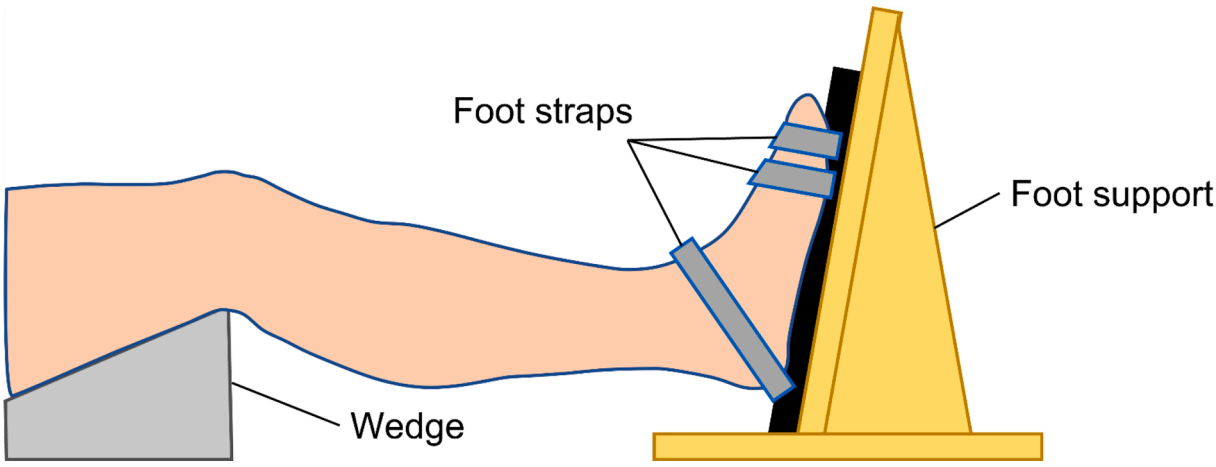

Fig. 1. Illustration of subject positioning for both 3DfUS and MRI measurements.

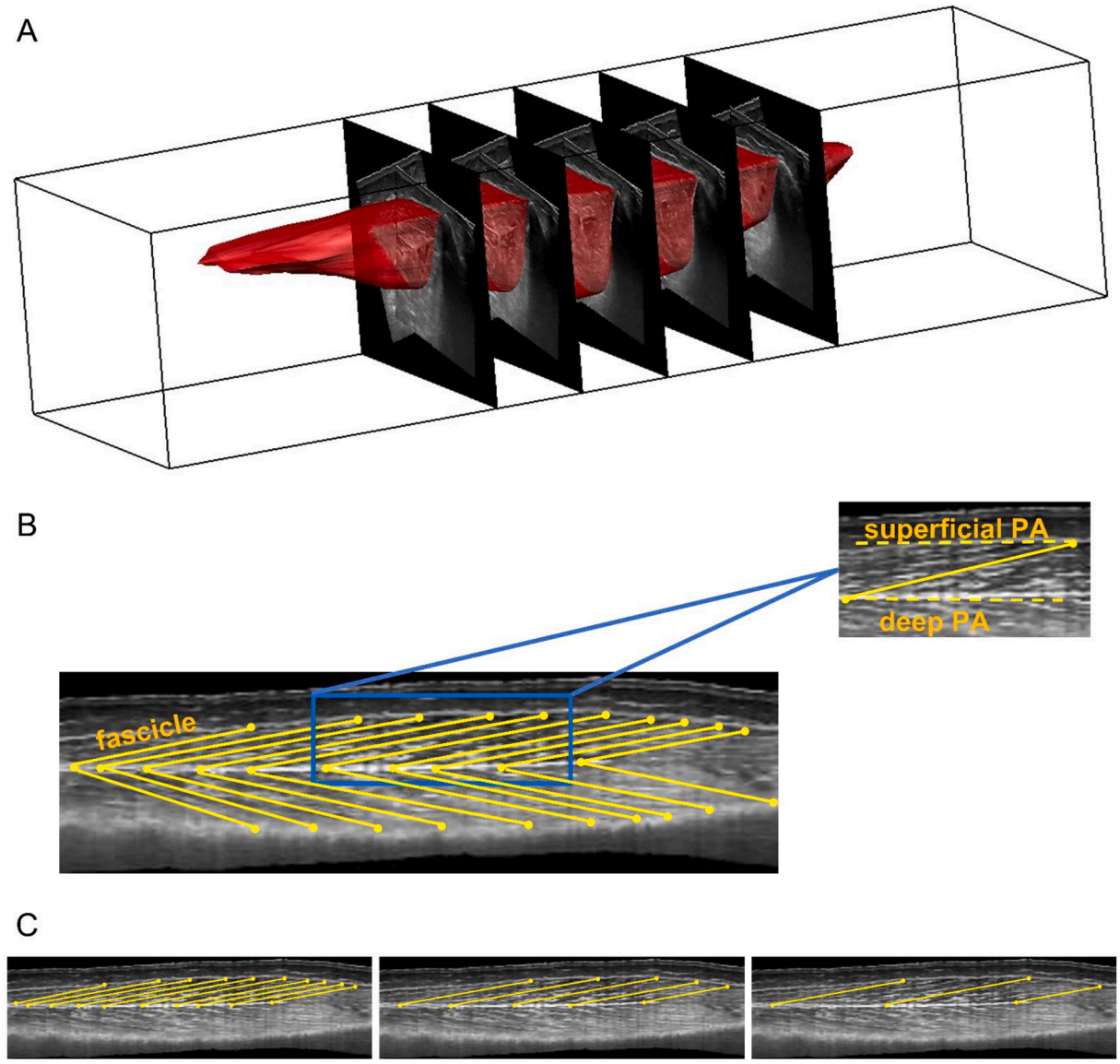

Fig. 2. (A) A sample of the reconstructed 3DfUS images of TA and the segmented surface model of TA. (B) Illustration of FL and PA measurement in 3DfUS of TA. Ten fascicles were manually selected in the mid-longitudinal plane where fascicles were clearly visible and intact. The FL of this muscle was defined as the mean length of these ten fascicles. Both superficial and deep PAs were measured. The PA of this muscle was defined as the mean value of all superficial and deep PAs. (C) Illustration of intra-rater reliability on the agreement of 3DfUS measurement with different number of digitized fascicles.

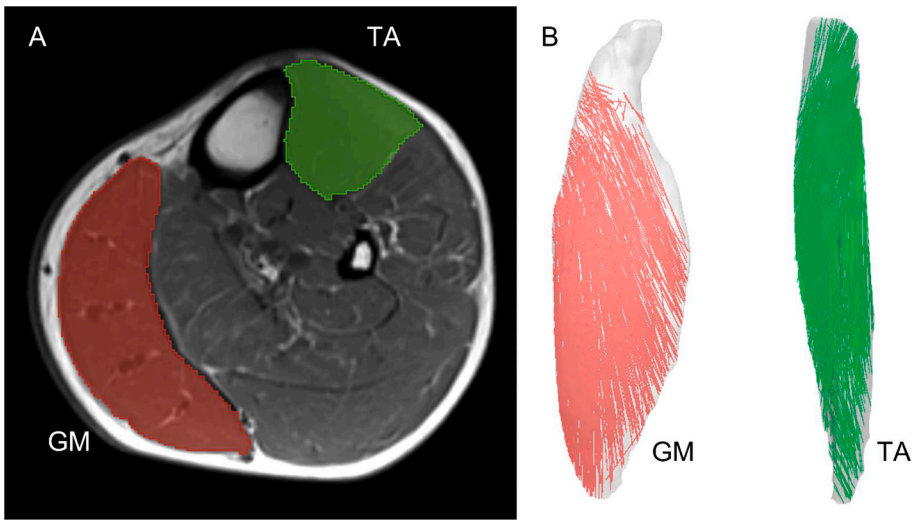

Fig. 3. Overview of muscle FL and PA measurement in MRI and DTI. (A) Muscle segmentation with the transverse view of one T1-weighted MR image. (B) 3D Muscle fibre reconstruction of GM and TA based on DTI data.

calculated directly from the surface model. Then, the tractography was performed in DSI studio (Yeh et al., 2013) to generate 1000 fibre tracts in the whole muscle, and each fibre tract was fitted with a 3D third-order polynomial curve. The curve was extended at each endpoint to the muscle surface along the slope of each endpoint. The FL was defined as the length of the extended fibre curve. The PA was defined as 90°minus the median angle between the fibre and the normal vectors of all meshes of the surface model within a search radius of 1.5 mm around the endpoint (Fig. 3B) (Bolsterlee et al., 2019). The reported FL was calculated as the mean length of 1000 fibres in each muscle, and PA was reported as the average of the deep and superficial PAs.

2.4. Data analysis

The mean, standard deviation (SD), and coefficient of variation (CV) of the 3DfUS measurement on muscle FL, PA, and volume were calculated. The intra-rater reliability of 3DfUS-based muscle morphological parameters quantification was evaluated in Group A, and the inter-session repeatability was evaluated in Group B. In particular, two aspects of the intra-rater reliability in FL and PA quantification were determined. The first aspect was the agreement of the manual post-processing procedures (selection of the fascicle plane and digitization of fascicles) performed two weeks apart by one processer. The second aspect was the agreement of 3DfUS measurement when different numbers of digitized fascicles (three, five, and ten digitized fascicles) were selected (Fig. 2**C**). The inter-session repeatability was defined as the agreement of 3DfUS measurement of muscle FL, PA, and volume for two different acquisitions. The intra-class correlation coefficient (ICC), standard error of measurement (SEM), and minimal detectable change (MDC) were used to evaluate the intra-rater reliability and inter-session repeatability of 3DfUS measurements (Weir, 2005). As a rule of thumb, the ICC above 0.9 was interpreted as excellent, and between 0.75 and 0.9 as good agreement (Koo and Li, 2016).

The mean differences between 3DfUS and MRI measurement on the muscle morphological parameters (FL, PA, and volume) were calculated. All differences were calculated as muscle parameters from 3DfUS minus MRI. The agreement of parameters identified by two methods was evaluated using the limits of agreement (LoA) approach (Bland and Altman, 1986). Furthermore, a paired *t*-test was used to investigate the differences of the muscle parameters identified by two methods. The statistical analysis was performed using SPSS 28.0 (SPSS Inc., Chicago, IL, USA). The significance level was set at 0.05.

3. Results

Good to excellent intra-rater reliability was observed in 3DfUS measurement on FL and PA in both TA and GM (Tables 1 and 2). Consistent estimations were found in both muscles when the same processor performed the manual post-processing at two different occasions. ICC were found higher than 0.90 and SEM were lower than 1.0 mm and 1.0° in FL and PA (Table 1). Digitizing different numbers of fascicles were also found having a negligible impact on the parameter estimation, where ICC were higher than 0.80 and SEM lower than 1.0 mm and 1.0° in both muscles (Table 2). The inter-session repeatability analysis revealed good to excellent agreement in FL, PA, and volume between two acquisitions (Table 3) with MDC lower than 1.0 mm, 1.0°, and 1.0 cm^3, respectively.

When using 3DfUS, the volumes of the TA and GM were underestimated by 3.7 cm^3 ($p = 0.179$) and 5.7 cm^3 ($p = 0.147$) compared to MRI measurement (Fig. 4 and Table 4). Furthermore, the FLs of the TA and GM were overestimated by 2.9 mm ($p = 0.054$) and 2.7 mm ($p = 0.054$), respectively (Fig. 5 and Table 4). These differences were not statistically significant. Compared to MRI-based measurement, a significantly larger PA was observed in the TA ($p = 0.019$) by 3DfUS, but no statistically significant difference was found in the GM ($p = 0.444$) (Fig. 6 and Table 4).

4. Discussion

This study evaluated the intra-rater reliability and inter-session repeatability of muscle FL, PA, and volume estimation using 3DfUS and compared them with parameters obtained using MRI in able-bodied subjects. 3DfUS measurements showed good to excellent intra-rater reliability and inter-session repeatability. Compared to MRI, no significant difference was observed in almost all muscle parameters in the TA and GM. Among the three parameters, the volume was the most accurately measured parameter using 3DfUS. To the best of our knowledge, this study is the first to compare 3DfUS and MRI measurements of muscle FL, PA, and volume in able-bodied subjects.

Most 3DfUS post-processing procedures are conducted manually and consist of several steps including muscle segmentation, fascicle plane selection, and fascicles digitization. To ensure the validity of the application, we evaluated the performance of 3DfUS with respect to intra-rater reliability and inter-session repeatability. Excellent inter-session repeatability was found in volume measurement, which was in line with a previous study of 3DfUS in a clinical application (Cenni et al.,

Table 1
Intra-rater reliability: the agreement of the manual post-processing procedures (selection of the fascicle plane and digitization of fascicles) performed two weeks apart by one processer.

Muscle	Parameter	Processing 1		Processing 2		ICC	SEM	MDC
		Mean ± SD	CV	Mean ± SD	CV			
TA	FL	41.5 ± 11.9 mm	0.29	40.2 ± 12.2 mm	0.30	0.99	0.2 mm	0.5 mm
	PA	15.2 ± 3.0°	0.20	15.2 ± 2.4°	0.16	0.90	0.3°	0.9°
GM	FL	41.6 ± 7.4 mm	0.18	39.8 ± 7.7 mm	0.19	0.93	0.8 mm	2.1 mm
	PA	18.7 ± 3.8°	0.20	17.2 ± 3.7°	0.21	0.90	0.3°	0.8°

SD: stand deviation, CV: coefficient of variation, ICC: intra-class correlation coefficient, SEM: standard error of measurement, MDC: minimal detectable change, TA: tibialis anterior, GM: gastrocnemius medialis, FL: fascicle length, PA: pennation angle.

Table 2
Intra-rater reliability: the agreement of 3DfUS measurement with different number of digitized fascicles.

Muscle	Parameter	Three fascicles		Five fascicles		Ten fascicles		ICC			SEM			MDC		
		Mean ± SD	CV	Mean ± SD	CV	Mean ± SD	CV	3 vs 5	3 vs 10	5 vs 10	3 vs 5	3 vs 10	5 vs 10	3 vs 5	3 vs 10	5 vs 10
TA	FL	40.7 ± 12.7 mm	0.31	40.9 ± 12.3 mm	0.30	41.5 ± 11.9 mm	0.29	0.99	0.99	0.98	0.1 mm	0.2 mm	0.3 mm	0.2 mm	0.6 mm	0.9 mm
	PA	14.8 ± 3.0°	0.20	15.1 ± 2.3°	0.15	15.2 ± 3.0°	0.20	0.92	0.99	0.89	0.3°	0.0°	0.4°	0.9°	0.1°	1.2°
GM	FL	39.7 ± 7.3 mm	0.18	39.8 ± 6.8 mm	0.17	41.6 ± 7.4 mm	0.18	0.99	0.96	0.94	0.1 mm	0.2 mm	0.4 mm	0.3 mm	0.5 mm	1.1 mm
	PA	18.0 ± 3.9°	0.21	18.7 ± 3.5°	0.19	18.7 ± 3.8°	0.20	0.96	0.98	0.96	0.2°	0.0°	0.2°	0.5°	0.1°	0.6°

SD: stand deviation, CV: coefficient of variation, ICC: intra-class correlation coefficient, SEM: standard error of measurement, MDC: minimal detectable change, TA: tibialis anterior, GM: gastrocnemius medialis, FL: fascicle length, PA: pennation angle.

Table 3
Inter-session repeatability: the agreement of 3DfUS measurement of muscle FL, PA, and volume for two different acquisition sessions.

Muscle	Parameter	Session 1		Session 2		ICC	SEM	MDC
		Mean ± SD	CV	Mean ± SD	CV			
TA	FL	39.0 ± 8.3 mm	0.21	37.7 ± 8.1 mm	0.21	0.99	0.0 mm	0.1 mm
	PA	15.4 ± 2.2°	0.14	13.9 ± 2.1°	0.15	0.81	0.1°	0.4°
	Volume	122.4 ± 30.6 cm^3	0.25	115.1 ± 31.8 cm^3	0.28	0.97	0.3 cm^3	0.8 cm^3
GM	FL	42.5 ± 5.5 mm	0.13	40.9 ± 5.4 mm	0.13	0.95	0.1 mm	0.2 mm
	PA	18.4 ± 2.4°	0.13	17.3 ± 2.5°	0.14	0.90	0.0°	0.1°
	Volume	226.3 ± 47.8 cm^3	0.21	218.7 ± 48.3 cm^3	0.22	0.99	0.1 cm^3	0.4 cm^3

SD: stand deviation, CV: coefficient of variation, ICC: intra-class correlation coefficient, SEM: standard error of measurement, MDC: minimal detectable change, TA: tibialis anterior, GM: gastrocnemius medialis, FL: fascicle length, PA: pennation angle.

2018a). The authors found excellent inter-session and inter-processer reliability in volume measurement for both children with spastic cerebral palsy and typically developed controls. Both studies have demonstrated the robustness of manual muscle segmentation. Fascicle plane selection is another step that may introduce errors in estimating FL and PA. By manually adjusting the tilting angle of the mid-longitudinal plane, a plane where the fascicles were most clearly visible was selected in our 3DfUS processing protocol. This procedure can be considered an analogy to tilting the ultrasound probe in 2D ultrasound measurement and was found repeatable indicated by the high ICC and low MDC values. In addition, digitizing different numbers of fascicles was found to have an ignorable influence on the FL and PA measurement, most likely because we digitized the fascicles located in mainly the muscle belly region. A previous study has reported that high variation in FLs existed between different regions within the same muscle according to DTI measurement (Aeles et al., 2022). However, due to the relatively poor ultrasound image quality in the distal and proximal ends of the muscle, we were not able to evaluate the regional FL differences within the same muscle using 3DfUS. Overall, our analysis demonstrated high intra-rater reliability and inter-session repeatability in 3DfUS measurement of muscle morphological parameters. These results indicated that 3DfUS could be a valuable tool for both cross-sectional and longitudinal assessment of muscle morphology in both healthy and pathological conditions.

3DfUS and MRI yielded highly comparable muscle FL and volume quantification in both the TA and GM. The results from LoA analysis showed that all the data points in the Bland-Altman plots were within the lower and upper LoA, suggesting no consistent bias of 3DfUS versus MRI measurement. Muscle volume was the most accurately measured parameter with a mean difference of −3.7 cm^3 and −5.7 cm^3 in the TA and GM. Muscle volume was regarded as a major determinant of joint torques (Fukunaga et al., 2001), and was correlated to maximum force-generating capacity (Knarr et al., 2013) and joint range of motion (Suga et al., 2021). MRI is regarded as a “gold standard” for *in vivo* measurement of muscle volume. Our finding was in line with an early study applying 3DfUS in muscle morphology measurements (Barber et al., 2009). Nevertheless, compared to MRI, we found muscle volume was slightly underestimated, which could have been due to the possibility of

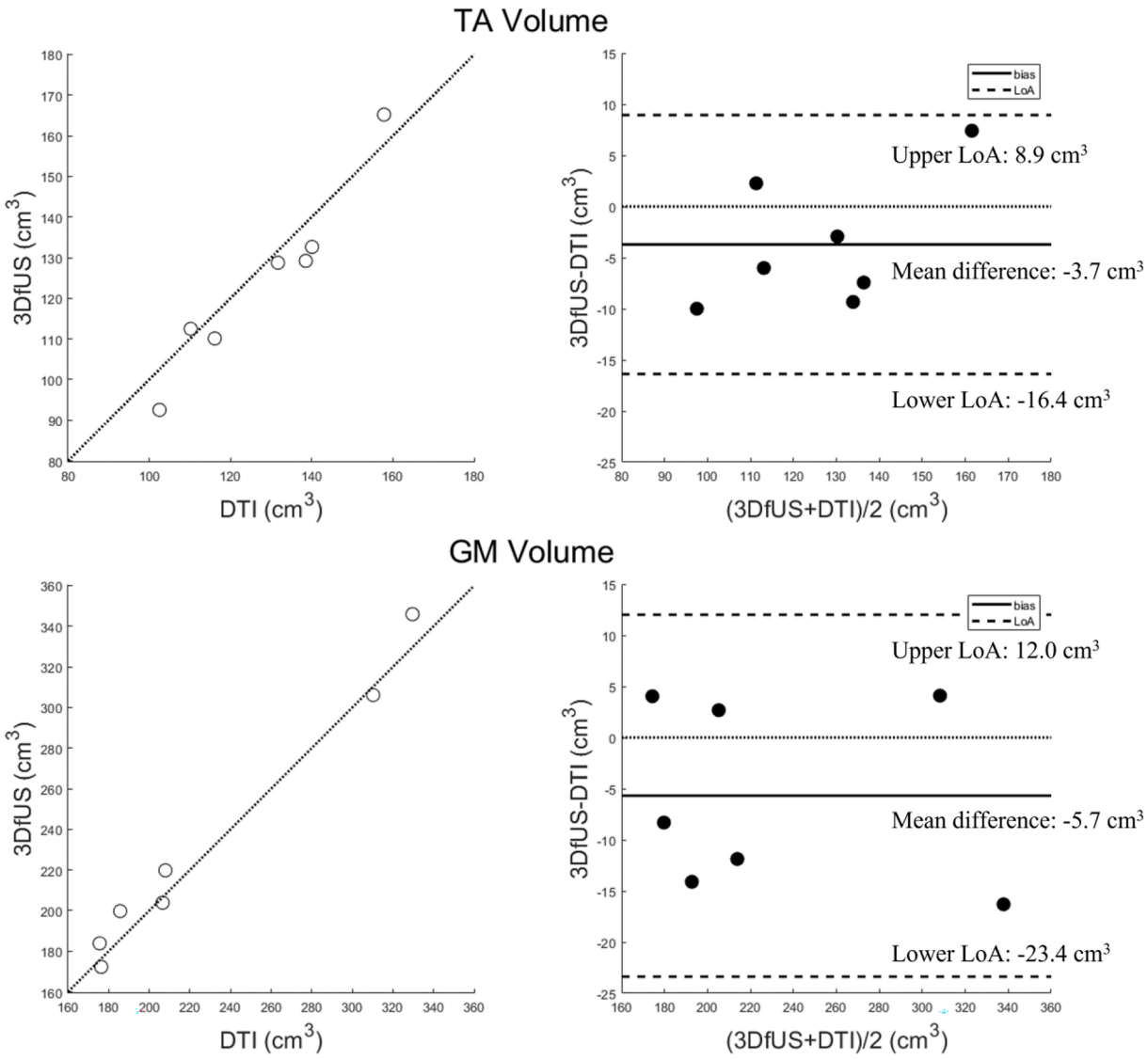

Fig. 4. Scatter plots of volume quantified by 3DfUS and DTI measurement in TA and GM (left column). In the scatter plots, the diagonal line illustrates the line of perfect agreement. Bland-Altman plots (right column) illustrate the difference in volume quantification by 3DfUS and DTI (3DfUS – DTI) versus mean measurement by two methods. The horizontal lines in the Bland-Altman plots correspond to mean difference and the upper and lower 95% limits of agreement.

Table 4
Comparison 3DfUS and MRI measurement of muscle morphological parameters (FL, PA, and volume) of TA and GM. The values were represented as mean ± standard deviation.

Muscle	Parameter	3DfUS	MRI	Difference	95% CI
TA	FL (mm)	41.5 ± 11.9	38.5 ± 13.7	2.9 ± 3.3	[−3.5, 9.3]
	PA (°) *	15.2 ± 3.0	13.3 ± 2.2	1.8 ± 1.5	[−1.2, 4.8]
	Volume (cm^3)	124.4 ± 22.9	128.1 ± 19.4	−3.7 ± 6.5	[−16.4, 8.9]
GM	FL (mm)	41.6 ± 7.4	39.0 ± 9.0	2.7 ± 3.0	[−3.1, 8.5]
	PA (°)	18.7 ± 3.8	18.4 ± 3.3	0.4 ± 1.1	[−1.9, 2.6]
	Volume (cm^3)	227.4 ± 64.8	233.1 ± 66.3	−5.7 ± 9.0	[−23.4, 12.0]

* $p < 0.05$.

the muscle not being fully covered in 3DfUS scanning. Apart from muscle volume, muscle FL and PA were also key factors influencing muscle force generation. The comparison of 3DfUS and DTI on muscle FL and PA estimation was therefore performed, though it should be kept in mind that there is no consensus on the "gold standard" of muscle FL and PA measurement *in vivo*. In this study, the mean difference in muscle FL estimation was less than 3.0 mm between 3DfUS and DTI (Table 2), which was similar to a previous comparison between 2D ultrasound and DTI. However, the variance of the FL difference between 2D ultrasound and DTI was reported more than 10 mm (Bolsterlee et al., 2015), suggesting that 3DfUS may be a more reliable method in muscle FL quantification. When measuring FL in 2D ultrasound, the ultrasound transducer should be oriented in an optimized position where the endpoints of the fascicle lie in the image and the transducer is perpendicular to the aponeurosis. Misalignment of the transducer could lead to errors in FL with more than 0.4 mm per degree of misalignment (Bolsterlee et al., 2016). Dual-transducer approaches were proposed to improve the limited FOV in 2D ultrasound (Brennan et al., 2017), but unequal compression of the two transducers on the skin might introduce extra errors. 3DfUS overcomes these limitations and requires no particular caution about the transducer's orientation, though additional time in training, data collection, and processing are needed to achieve a reliable 3DfUS measurement. Regarding PA measurement, although significant between-methods differences were found in PA of the TA, the discrepancy was small (1.84°). No direct comparison study between DTI and

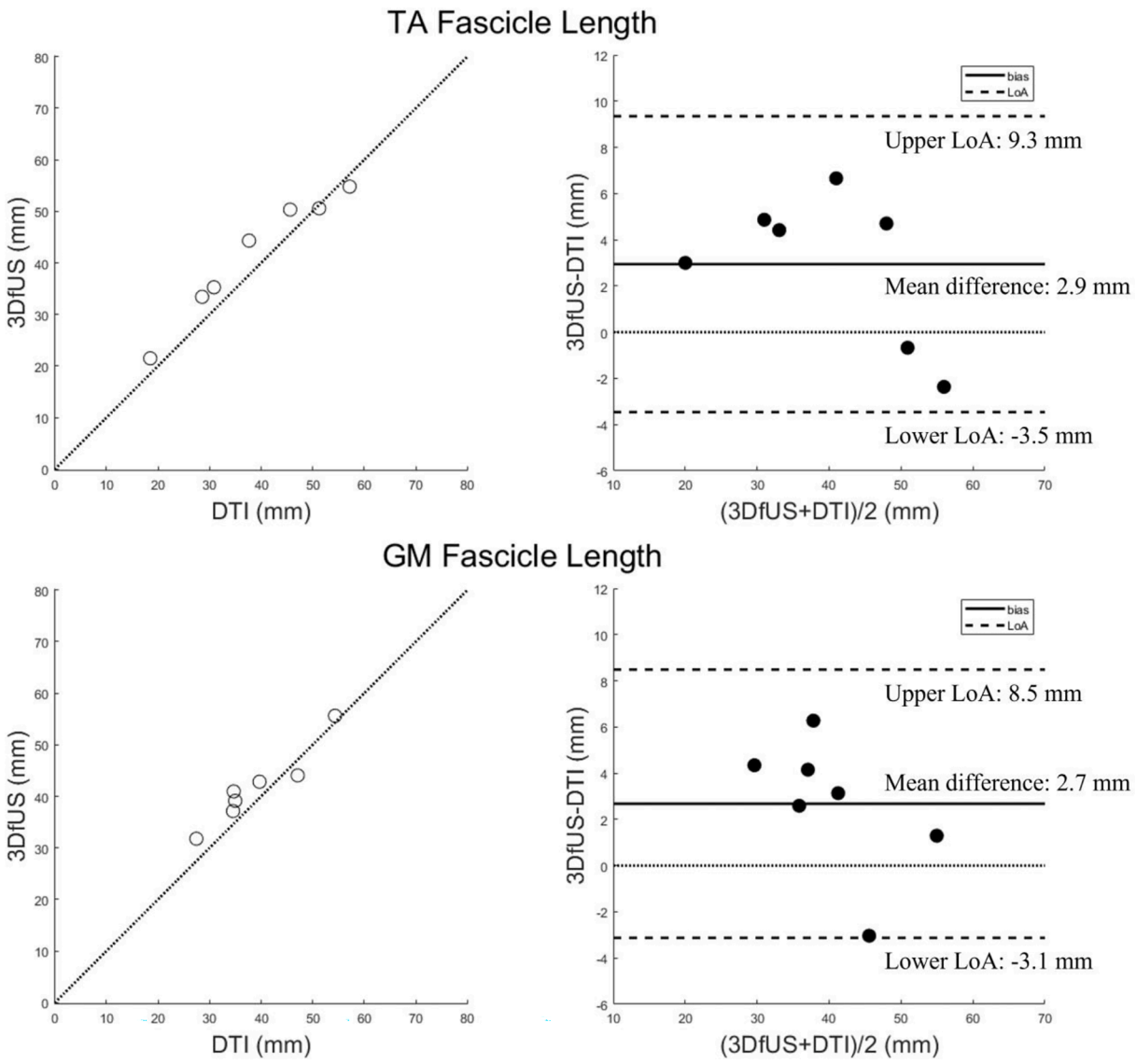

Fig. 5. Scatter plots of fascicle length quantified by 3DfUS and DTI measurement in TA and GM (left column). In the scatter plots, the diagonal line illustrates the line of perfect agreement. Bland-Altman plots (right column) illustrate the difference in fascicle length quantification by 3DfUS and DTI (3DfUS - DTI) versus mean measurement by two methods. The horizontal lines in the Bland-Altman plots correspond to mean difference and the upper and lower 95% limits of agreement.

3DfUS measurement in able-bodied subjects was found. In our previous study, significantly different PA was found in the TA when comparing 2D ultrasound and DTI in post-stroke individuals (median difference 25.4°) (Korting et al., 2019). There were several reasons for this, such as muscle deformation caused by transducer compression and rheological property changes in the spastic muscle after stroke. Due to the unclear between-compartment border in the MRI, we did not divide TA into superficial and deep compartments. Further studies analyzing the two sub-compartments separately in 3DfUS may improve the PA measurement. Overall, compared to 2D ultrasound, 3DfUS provided a more precise and reliable approach to quantifying muscle FL, PA, and volume *in vivo*.

There were several limitations in this study. The sample size of this study was small (7 and 9 in group A and B) due to the challenges in subjects' recruitment and the extensive measurement protocol. During fascicle digitalization in 3DfUS, it was challenging to find a proper mid-longitudinal plane that visualizes all the intact fascicles passing through the same plane. FL was calculated as the Euclidean distance between the two endpoints in 3DfUS. This is still an acceptable approximation in passive conditions, but fascicles should be outlined as curves. Further developments on tracking the fascicles as curves in 3DfUS images are needed. An automatic fascicle tracking approach using deep learning was proposed for 2D ultrasound (Cronin et al., 2020), which could potentially be expanded to track fascicles in 3DfUS images. In addition, the muscle architecture parameters might change due to the deformation induced by transducer compression, even though a generous amount of acoustic gel was used. To minimize muscle deformation, a gel pad was proposed to place between the transducer and skin, reducing muscle deformation by 46% in typically developed children (Cenni et al., 2018b). However, this approach sacrifices the imaging depth, which limited its broader application in adults, especially in some populations with more subcutaneous fat.

In conclusion, this study demonstrated that 3DfUS can reliably quantify muscle FL, PA, and volume in able-bodied subjects. 3DfUS measurement achieved good to excellent intra-rater reliability and inter-session repeatability. Compared to MRI measurement, 3DfUS had errors within 2.9 mm for FL, 1.8° for PA, and 5.7 cm^3 for volume. Compared to other imaging modalities, ultrasound imaging has the advantages of low-cost, good accessibility, and portability. Thus, researchers may consider 3DfUS as an alternative to MRI for muscle morphological

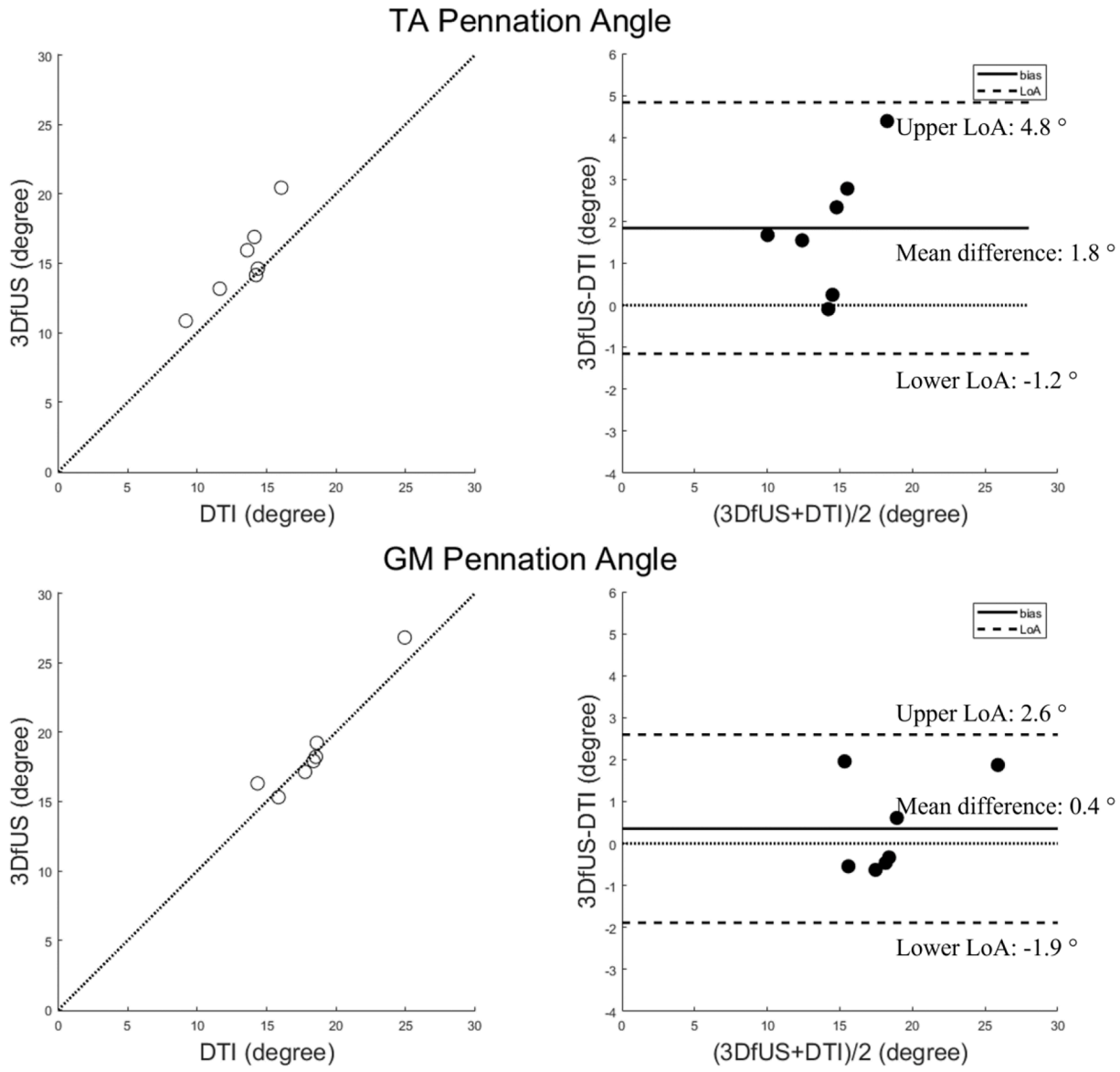

Fig. 6. Scatter plots of pennation angle quantified by 3DfUS and DTI measurement in TA and GM (left column). In the scatter plots, the diagonal line illustrates the line of perfect agreement. Bland-Altman plots (right column) illustrate the difference in pennation angle quantification by 3DfUS and DTI (3DfUS - DTI) versus mean measurement by two methods. The horizontal lines in the Bland-Altman plots correspond to mean difference and the upper and lower 95% limits of agreement.

architecture evaluation.

CRediT authorship contribution statement

Zhongzheng Wang: Visualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Antea Destro:** Investigation, Formal analysis, Writing - review & editing. **Sven Petersson:** Investigation, Writing - review & editing. **Francesco Cenni:** Methodology, Writing - review & editing. **Ruoli Wang:** Supervision, Investigation, Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Zhongzheng Wang thanks the financial support from Chinese Scholarship Council and Promobilia Foundation (Ref. 19088 & Ref. 21033). Francesco Cenni is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 101028724.

References

Aeles, J., Bolsterlee, B., Kelp, N.Y., Dick, T.J.M., Hug, F., 2022. Regional variation in lateral and medial gastrocnemius muscle fibre lengths obtained from diffusion tensor imaging. J. Anat. 240 (1), 131–144.

Barber, L., Barrett, R., Lichtwark, G., 2009. Validation of a freehand 3D ultrasound system for morphological measures of the medial gastrocnemius muscle. J. Biomech. 42 (9), 1313–1319.

Bland, J.M., Altman, D., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327 (8476), 307–310.

Bolsterlee, B., Veeger, H.E., van der Helm, F.C., Gandevia, S.C., Herbert, R.D., 2015. Comparison of measurements of medial gastrocnemius architectural parameters from ultrasound and diffusion tensor images. J. Biomech. 48 (6), 1133–1140.

Bolsterlee, B., Gandevia, S.C., Herbert, R.D., 2016. Effect of transducer orientation on errors in ultrasound image-based measurements of human medial gastrocnemius muscle fascicle length and pennation. PLoS One 11 (6), e0157273.
Bolsterlee, B., D'Souza, A., Herbert, R.D., 2019. Reliability and robustness of muscle architecture measurements obtained using diffusion tensor imaging with anatomically constrained tractography. J. Biomech. 86, 71–78.
Brennan, S.F., Cresswell, A.G., Farris, D.J., Lichtwark, G.A., 2017. In vivo fascicle length measurements via B-mode ultrasound imaging with single vs dual transducer arrangements. J. Biomech. 64, 240–244.
Cenni, F., Monari, D., Desloovere, K., Aertbelien, E., Schless, S.H., Bruyninckx, H., 2016. The reliability and validity of a clinical 3D freehand ultrasound system. Comput. Methods Programs Biomed. 136, 179–187.
Cenni, F., Schless, S.H., Bar-On, L., Aertbeliën, E., Bruyninckx, H., Hanssen, B., Desloovere, K., 2018a. Reliability of a clinical 3D freehand ultrasound technique: analyses on healthy and pathological muscles. Comput. Methods Programs Biomed. 156, 97–103.
Cenni, F., Schless, S.H., Monari, D., Bar-On, L., Aertbeliën, E., Bruyninckx, H., Hanssen, B., Desloovere, K., 2018b. An innovative solution to reduce muscle deformation during ultrasonography data collection. J. Biomech. 77, 194–200.
Charles, J.P., Grant, B., D'Aout, K., Bates, K.T., 2020. Subject-specific muscle properties from diffusion tensor imaging significantly improve the accuracy of musculoskeletal models. J. Anat. 237 (5), 941–959.
Cronin, N. J., Finni, T., Seynnes, O., 2020. Fully automated analysis of muscle architecture from B-mode ultrasound images with deep learning. arXiv preprint arXiv:2009.04790.
Damon, B.M., Ding, Z., Anderson, A.W., Freyer, A.S., Gore, J.C., 2002. Validation of diffusion tensor MRI-based muscle fiber tracking. Magn. Reson. Med. 48 (1), 97–104.
De Oliveira, L.F., Menegaldo, L.L., 2010. Individual-specific muscle maximum force estimation using ultrasound for ankle joint torque prediction using an EMG-driven Hill-type model. J. Biomech. 43 (14), 2816–2821.
Delabastita, T., Afschrift, M., Vanwanseele, B., De Groote, F., 2020. Ultrasound-based optimal parameter estimation improves assessment of calf muscle-tendon interaction during walking. Ann. Biomed. Eng. 48 (2), 722–733.
D'Souza, A., Bolsterlee, B., Lancaster, A., Herbert, R.D., 2019. Muscle architecture in children with cerebral palsy and ankle contractures: an investigation using diffusion tensor imaging. Clin. Biomech. 68, 205–211.
Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R., 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn. Reson. Imag. 30 (9), 1323–1341.
Fukunaga, T., Miyatani, M., Tachi, M., Kouzaki, M., Kawakami, Y., Kanehisa, H., 2001. Muscle volume is a major determinant of joint torque in humans. Acta Physiol. Scand. 172 (4), 249–255.
Hanssen, B., De Beukelaer, N., Schless, S.H., Cenni, F., Bar-On, L., Peeters, N., Molenaers, G., Van Campenhout, A., Van den Broeck, C., Desloovere, K., 2021. Reliability of processing 3-D freehand ultrasound data to define muscle volume and echo-intensity in pediatric lower limb muscles with typical development or with spasticity. Ultrasound Med. Biol. 47 (9), 2702–2712.
Holzbaur, K.R., Murray, W.M., Gold, G.E., Delp, S.L., 2007. Upper limb muscle volumes in adult subjects. J. Biomech. 40 (4), 742–749.
Klimstra, M., Dowling, J., Durkin, J.L., MacDonald, M., 2007. The effect of ultrasound probe orientation on muscle architecture measurement. J. Electromyogr. Kinesiol. 17 (4), 504–514.
Knarr, B.A., Ramsay, J.W., Buchanan, T.S., Higginson, J.S., Binder-Macleod, S.A., 2013. Muscle volume as a predictor of maximum force generating ability in the plantar flexors post-stroke. Muscle Nerv. 48 (6), 971–976.
Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15 (2), 155–163.
Korting, C., Schlippe, M., Petersson, S., Pennati, G.V., Tarassova, O., Arndt, A., Finni, T., Zhao, K., Wang, R., 2019. In vivo muscle morphology comparison in post-stroke survivors using ultrasonography and diffusion tensor imaging. Sci. Rep. 9 (1), 11836.
Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. J. Magn. Reson. Imag. 13 (4), 534–546.
Mitsiopoulos, N., Baumgartner, R., Heymsfield, S., Lyons, W., Gallagher, D., Ross, R., 1998. Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. J. Appl. Physiol. 85 (1), 115–122.
Nelson, C.M., Murray, W.M., Dewald, J.P., 2018. Motor impairment-related alterations in biceps and triceps brachii fascicle lengths in chronic hemiparetic stroke. Neurorehabil. Neural Repair 32 (9), 799–809.
Oudeman, J., Mazzoli, V., Marra, M.A., Nicolay, K., Maas, M., Verdonschot, N., Sprengers, A.M., Nederveen, A.J., Strijkers, G.J., Froeling, M., 2016. A novel diffusion-tensor MRI approach for skeletal muscle fascicle length measurements. Physiol. Rep. 4 (24), e13012.
Padhiar, N., Al-Sayeagh, H., Chan, O., King, J., Maffulli, N., 2008. Pennation angle of the soleus in patients with unilateral Achilles tendinopathy. Disabil. Rehabil. 30 (20–22), 1640–1645.
Prager, R., Gee, A., Treece, G., Berman, L., 2002. Freehand 3D ultrasound without voxels: volume measurement and visualisation using the Stradx system. Ultrasonics 40 (1–8), 109–115.
Rana, M., Wakeling, J.M., 2011. In-vivo determination of 3D muscle architecture of human muscle using free hand ultrasound. J. Biomech. 44 (11), 2129–2135.
Rehmann, R., Froeling, M., Rohm, M., Forsting, J., Kley, R.A., Schmidt-Wilcke, T., Karabul, N., Meyer-Friessem, C.H., Vollert, J., Tegenthoff, M., Vorgerd, M., Schlaffke, L., 2020. Diffusion tensor imaging reveals changes in non-fat infiltrated muscles in late onset Pompe disease. Muscle Nerve 62 (4), 541–549.
Reid, S.L., Pitcher, C.A., Williams, S.A., Licari, M.K., Valentine, J.P., Shipman, P.J., Elliott, C.M., 2015. Does muscle size matter? The relationship between muscle size and strength in children with cerebral palsy. Disabil. Rehabil. 37 (7), 579–584.
Suga, T., Terada, M., Tomoo, K., Miyake, Y., Tanaka, T., Ueno, H., Nagano, A., Isaka, T., 2021. Association between plantar flexor muscle volume and dorsiflexion flexibility in healthy young males: ultrasonography and magnetic resonance imaging studies. BMC Sports Sci. Med. Rehabil. 13 (1), 1–8.
Takahashi, K., Shiotani, H., Evangelidis, P.E., Sado, N., Kawakami, Y., 2022. Three-dimensional architecture of human medial gastrocnemius fascicles in vivo: regional variation and its dependence on muscle size. J. Anat. 241 (6), 1324–1335.
Weide, G., van der Zwaard, S., Huijing, P.A., Jaspers, R.T., Harlaar, J., 2017. 3D ultrasound imaging: fast and cost-effective morphometry of musculoskeletal tissue. J. Vis. Exp. 129.
Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Cond. Res. 19 (1), 231–240.
Yeh, F.-C., Verstynen, T.D., Wang, Y., Fernández-Miranda, J.C., Tseng, W.-Y.-I., 2013. Deterministic diffusion fiber tracking improved by quantitative anisotropy. PLoS One 8 (11), e80713.