# A comparison of 3D ultrasound to MRI for the measurement and estimation of gastrocnemius muscle volume in adults and young people with and without cerebral palsy

Noorkoiv, M., Theis, Nicola and Lavelle, G.

## Abstract

### Introduction

Muscle volume (MV) is an important parameter for understanding muscle morphology and adaptations to training, growth and pathology. In this study, we assessed the validity of freehand 3D ultrasound (3DUS) for measuring medial gastrocnemius MV in adults, typically developing children (TD) and children with cerebral palsy (CP). We also assessed the validity between our direct measures of MV and estimates derived from anatomical cross sectional area (ACSA) and muscle length (ML), using previously outlined methods.

### Materials and Methods

The medial gastrocnemius of all groups was scanned with 3DUS and MRI. Images from both methods were digitised to derive MV, ACSA and ML. Measured MV was compared between methods and compared to estimated MV derived from recently published algorithms.

### Results

MV had a mean difference of -0.13% (standard error of estimate (SEE)=2.23%, $R^2$=0.99) between MRI and 3DUS and 19.82% (SEE=4.73% and $R^2$=0.99) and -3.11% (SEE=6.55%, $R^2$=0.99) mean differences between the measured and estimated MV from two methods of estimation.

### Conclusions

The 3DUS is a valid method for the measurement of MV in adults, TD children and those with CP. Estimation methods of MV may be useful in clinical practise, but require further replication on various populations and careful methodological consideration.


**Keywords:** muscle morphology; muscle size; freehand 3D ultrasound; muscle imaging

## Introduction

Skeletal muscle size is an important determinant of force-generation and physical function (Fukunaga et al., 2001; O'Brien et al., 2009). The estimation of MV and ML in particular, are often measures of interest in intervention or cross-sectional studies to examine muscle hypertrophy (Roig et al., 2008), atrophy (Aagaard et al., 2010), or adaptations to neuromuscular pathologies (Fry et al., 2007). The 'gold standard' method for the assessment of muscle size is magnetic resonance imaging (MRI), which provides 3D high-resolution *in vivo* images of MV and ML (Mitsiopoulos et al., 1998). However, the use of MRI is an expensive technique with strong magnetic fields limiting the assessment of individuals with metal implants or epilepsy. Additionally, MRI requires long scanning times (15-20 mins for calf muscles), which require participants to remain still for the duration of the scan a limitation, which may become particularly apparent in clinical populations (e.g. CP, osteoarthritis) and children. An alternative measure of *in vivo* muscle size can be obtained using 3DUS. For this purpose, a series of 2D cross-sectional ultrasound images are simultaneously recorded whilst the position and orientation of the ultrasound probe are tracked using 3D motion analysis (http://mi.eng.cam.ac.uk/~rwp/stradwin/). A reconstruction of the 2D ultrasound images in 3D space is then derived to give estimates of muscle morphology. This method may be favourable compared to MRI due its lower cost, portability and considerably shorter scanning times.

Freehand 3DUS has been used to provide reliable measures of MV and ML in both in *vitro* (Weide et al., 2017) and in *vivo* studies of TD children and those with CP (Schless et al., 2017). Dissected medial gastrocnemius MV and fascicle length from 4 cadavers showed significantly high correlations with MVs and fascicle lengths measured with

3DUS using a water submersion method and callipers (Weide et al., 2017). Using a similar method, Barber et al. (2009) found that freehand 3DUS in healthy young adults had very good agreement with MRI, overestimating medial gastrocnemius MV by only 1.1 ± 3.8% and underestimating ML by 1.3% (± 2.2). However, there may be two main reasons why this method cannot be assumed to produce the same valid results for other populations, such as TD children and those with CP, despite the increasing wide spread use of freehand 3DUS in such studies. First, the assessment of calf MV with freehand 3DUS has only been compared to MRI *in vivo* when participants were knelt down in a water bath. This method, which requires a constant joint position, would be inappropriate in those with CP and thus, an alternative method to enhance visualization and allow maximal contact between the ultrasound probe and the curved surface of the muscle would be needed. We propose using a custom shaped gel pad as a more alternative to a water bath. Second, digitisation of the ACSA may be challenging in hyperechoic CP muscles (Pitcher et al., 2015), where segmentation of the muscle border may be more difficult. Thus, the first purpose of the study was to assess the validity and intra-digitiser reliability of 3DUS compared to MRI in adults, TD children and children with CP.

In order to develop a simple, clinically feasible method of MV quantification, various methods of *estimating* MV from simpler imaging techniques have been proposed (Albracht et al., 2008; Mersmann et al., 2014). As MV cannot be measured using 2D ultrasound, surrogate outcome parameters such as ACSA are often used (Schless et al., 2017; Vanmechelena et al., 2018). Vanmechelena et al. (2018) demonstrated that medial gastrocnemius MV could be reliably estimated in a clinically feasible manner in TD children and those with CP, from measures of maximal ACSA ($ACSA_{max}$) and ML measured from MRI. However, measures of whole anatomical muscle length

($ML_{anatomical}$) without the use of MRI is challenging, and finding the $ACSA_{max}$ can also take considerable time. A simpler method of estimation was proposed by Schless et al. (2017), which used 3DUS to determine ACSA at 50% ML. However, this method used an unconventional definition of the proximal origin of the medial gastrocnemius at the tibia, which will exclude the most proximal end of the estimated muscle volume. Therefore, the second purpose of the study was to assess the validity of using 3DUS to derive MV using the estimation methods outlined by Schless et al. (2017) and Vanmechelena et al. (2018). Estimating MV from 3DUS data would save digitising time and thus, be less time-consuming than measuring MV directly. Thus, the third aim of this study was to explore the practical aspects of feasibility in relation to measurement accuracy of directly measuring MV vs estimating MV with two recently proposed simpler estimation methods.

## Materials and Methods

*Participants*

Six adults (four male, two female, mean ± SD) age = 29.8 ± 9.0 y, body mass = 72.3 ± 19.7 kg, height = 171.7 ± 0.1 cm), six TD adolescents (one male , four female, 12.2 ± 3.5 y, body mass = 42.5 ± 16.5 kg, height = 148.8 ± 0.1 cm) and six adolescents with diplegic or hemiplegic CP, Gross Motor Function Classification System (GMFCS) level I or II (four male, two female, age = 15.0 ± 3.1 y, body mass = 50.2 ± 8.6 kg, height = 151.9 ± 0.1 cm) participated in this study. Participants in all groups had no recent or recurrent Achilles tendon or lower leg musculoskeletal injury and/or surgery, no muscle spasticity or tremor, no metal implants or mental impairment that would limit participation in the study. Participants were asked to avoid strenuous physical activity 48 hours prior the testing session. All participants were asked to complete a medical history questionnaire, MRI safety checklist and provide written informed consent. For

the participants under 16 years, parents or legal guardians also provided written informed consent. The study was approved by the local University ethics committee.

*Experimental design*

Participants attended the laboratory on one occasion. Prior to ultrasound and MRI scanning, participants lay prone on an examination table with the foot extended off the edge of the table for 10 minutes, to allow fluid shifts in the muscle to stabilise (Cerniglia et al., 2007). Two scans of the medial gastrocnemius muscle were acquired using a 3DUS method. Following this, the participant was transferred to the MRI table and a contiguous MRI scan of the dominant lower leg for adults and TD adolescents and more affected leg of adolescents with CP was acquired. During 3DUS and MRI imaging procedures, the participants were required to remain still and relaxed, with the hip and knee joint fully extended and the ankle angle in a supported resting position.

*Freehand 3DUS setup and calibration*

Two-dimensional B-mode ultrasound images of the transverse plane of the medial gastrocnemius muscle combined with 3D motion data were collected using Stradwin v5.1 software (Mechanical Engineering, Cambridge University, Cambridge, UK; http://mi.eng.cam.ac.uk/~rwp/stradwin (Treece, 2003)). B-mode ultrasound images were recorded at 40 Hz using a PC-based ultrasound scanner and a 10 MHz linear transducer with 60 mm field of view (HL9.0/60/128Z-2, Telemed LogicScan 128 EXT-1Z system, Lithuania). Position and orientation of the ultrasound probe were recorded by the Optitrack V120:Trio (NaturalPoint, Inc., Oregon, USA) optical position sensor by tracking four markers rigidly attached to the probe. Before data acquisition, the 3D freehand system setup quality control test was done on imaging temporal and spatial parametersby following the single-wall phantom calibration guidelines provided by

Stradwin software (Prager RW, 1998). The temporal calibration (the relative lag between the position and image streams) was estimated by matching the image motion with the readings from the position sensor. The result were presented in terms of the lag, and also a confidence value, representing how well the image and position streams correlated. Few results with confidences exceeding 95% were obtained for temporal calibration until Stradwin displayed a temporal calibration result that was in line with the> 95% confidence consensus. The spatial calibration was performed using a planar surface, i.e. the floor of a flat-bottomed water bath that was clearly visible in the ultrasound image (the phantom plane), which was scanned and a least square fit was performed to estimate the three best translations and rotations of the line data that was on the floor of a flat-bottomed water bath, and fitted the plane. Finally, the system was set to an image depth of 50 mm.

*Freehand 3DUSd imaging and analysis*

One to three sweeps of the medial gastrocnemius muscle was performed depending on the size of the muscle. Sweeps were obtained by moving the probe across the calf starting just above the line connecting the most superficial aspect of the lateral femoral condyle to the distal musculotendinous junction. The speed of the probe movement was approximately 5 cm/s. A consistent, minimal pressure was applied with the probe to avoid compression of the muscle. This was aided by the application of a transmission gel to improve acoustic coupling. In order to keep the pressure minimal and avoid gaps at the lateral edges of the image a customised curved gel pad was held in between the skin and the probe’s imaging surface.

The medial gastrocnemius muscle boundaries were manually digitised in all corresponding axial plane images using digitising and volume reconstruction functions

in Stradwin software (v5.1 software (Mechanical Engineering, Cambridge University, Cambridge, UK). Although minimal probe pressure during the scanning procedure was assured, the position and pressure correction functions were applied to MV data prior to digitisation to increase the clarity of reslices and precision of the data, as suggested by Treece et al. (2002). Every third image frame (3 mm) of the muscle sweep was digitised (no inter-frame gaps) and reconstructed into a rendered 3D muscle. For data with multiple sweeps, Stradwin divided the space into a number of partitions, with each partition associated with a particular sweep. After the digitisation procedure, the image reslice window was adjusted to be perpendicular to the longitudinal axis of the medial gastrocnemius and the segmented data was exported as voxels (resolution = 1 pixel width, spacing = 1 mm). The 3DUS image acquisition and analysis was practiced by the research physiotherapist on 15 participants prior to undertaking this study (approximately 35 hours of training and practicing).

*MRI imaging and analysis*

Participants lay supine on the MRI scanning table with right foot in 10° plantarflexion. This was measured with a goniometer, and a wedge was placed under the foot to support and maintain the plantar flexed position throughout the scan. Additional sand bags were placed on both sides of the foot to avoid lateral rotation. Axial plane contiguous T1-weighted MRI images were acquired along the length of the magnet bore for the entire length of the medial gastrocnemius muscle from above the femoral condyles to below the calcaneus. The participants were scanned with a 3T whole body MRI system with total imaging technology system (MAGNETOM Trio Syngo MR 2004A, Siemens, Erlangen, Germany) using a gradient echo sequence with the following parameters: time to echo = 8.3 ms, time to repetition = 500 ms, field of view = 610 x 406 mm, matrix = 320 x 256 pixels, 4 mm slice thickness and 0 mm of inter-

slice gap. Images were acquired of the same leg as with the 3DUS. The medial gastrocnemius muscle boundaries (every 4-mm slice with no gaps) were digitised in Osirix software (OsiriXv.4.0; Pixmeo, Switzerland), each muscle was digitised twice by an experienced digitiser and the average of two MV values was used for analysis.

*Data analysis*

For 3DUS, medial gastrocnemius ML was measured as the distance between the inferior margin of the medial condyle of the tibia to the distal musculotendinous junction, in the sagittal plane (Fig. 1). The inferior margin of the medial condyle of the tibia was chosen to represent the proximal end of the muscle because the muscle boundaries at the true origin (medial condyle of the femur) were not clear enough in all participants. The average of two repeated digitised MV values was used for final analyses.

For comparison to 3DUS and the estimation methods outlined by Schless et al. (2017) and Vanmechelena et al. (2018), muscle volume was calculated from two different muscle length and ACSA measures. First, ML was measured from the distal insertion of the medial gastrocnemius to the inferior margin of the medial condyle of the tibia ($ML_{MCT}$). The $ACSA_{max}$ was measured as a cross-section, perpendicular to the longitudinal axis of the participant's body with respect to the scanner table and expressed as a percentage distance from the proximal origin of the medial gastrocnemius muscle. This allowed the comparison of MVs assessed with MRI and 3DUS.

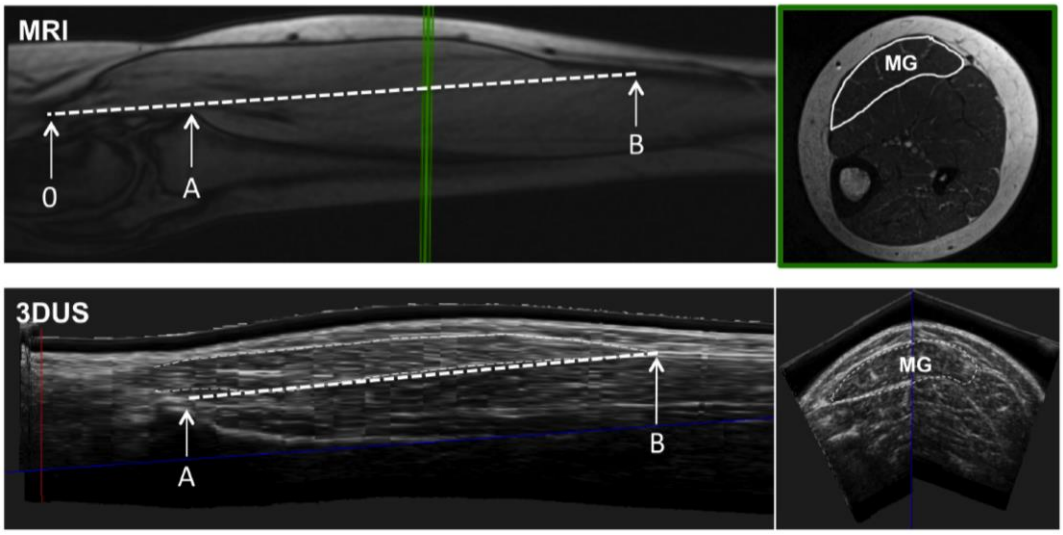


Figure 1. The assessment of muscle length with MRI and 3DUS methods. A sagittal view of the shank is shown to illustrate the muscle length assessment. 0 = anatomical medial gastrocnemius muscle origin, A = inferior margin of the medial condyle of the tibia, B = medial gastrocnemius muscle insertion at the musculotendinous junction, MG = medial gastrocnemius. Total muscle length from its anatomical origin to insertion is shown as a line from point 0 to point B on the MRI sagittal image. The lines connecting points A and B on both MRI and 3DUS sagittal images shows how muscle length was assessed for the 3DUS volume validity and comparison to $Est_1$.

Second, ML from the distal insertion of the medial gastrocnemius to the inferior margin of the medial condyle of the tibia ($ML_{MCT}$) was combined with ACSA at 50% of the distance from the inferior margin of the medial condyle of the tibia ($ACSA_{50\%}$). This allowed a comparison of MV measured with 3DUS, to the volume estimation outlined by Schless et al. (2017) (Fig. 1) which is referred in the present paper as $MV_{est1}$:

$$MV_{est1} = 5.472+0.526\ (ACSA_{50\%} \times ML_{MCT}) \quad (2)$$

Finally, medial gastrocnemius length was calculated from the proximal origin at the medial condyle of the femur to its distal insertion ($ML_{anatomical}$), and combined with

$ACSA_{max}$. This allowed a comparison of MV measured with MRI to the volume estimation outlined by Vanmechelena et al. (2018) (Fig. 1) and is referred in the present paper as $MV_{est2}$:

$$MV_{est2} = \left((ACSA_{max} \times ML_{anatomical}) - Offset\right) \times FF \quad (1)$$

where, FF = form factor, 0.619, Offset = 0 (Vanmechelena et al., 2018).

Total scanning time (including the identification of the muscle borderlines and imaging data saving), for a single medial gastrocnemius MV assessment with a 3DUS method by an experienced sonographer ranged between 1 to 2 minutes depending on the size of the muscle. Two successful scans were taken for each participant. Ultrasound image processing, segmentation and reconstruction time for a single MV calculation ranged between 15 to 25 min, giving a total time for MV assessment of about 17-30 min. For MRI method, the scanning time including participant positioning ranged between 15 to 20 min, depending on leg length. The MRI image processing, segmentation and reconstruction time for a single medial gastrocnemius muscle ranged between 10 to 20 min, giving a total maximal time of MV assessment of approximately 30-40 min.

*Statistical analysis*

A series of one-way ANOVAs were used to determine differences in MV, ML, $ACSA_{max}$, $ACSA_{max}$ location and $ACSA_{50\%}$ in adults, TD children and children with CP. Estimation of $MV_{est1}$ was compared to MV obtained with using 3DUS. Estimation of $MV_{est2}$ was compared to MV obtained with using MRI. Bland-Altman analysis was performed to compare MV measured with MRI and 3DUS. Linear regression analyses between

measured MV and $MV_{est1}$ and $MV_{est2}$ were used to calculate the $R^2$ and standard error of estimate (SEE) values. The SEE was also reported as a percentage of the average MV. A series of paired samples t-tests were used to compare ML, $ACSA_{max}$, location of $ACSA_{max}$ and $ACSA_{50\%}$ when assessed with MRI vs 3DUS. Finally, a one-way MANOVA was used to test whether there are any differences between groups (3 groups) on MV assessed with MRI compared to the $MV_{est1}$ and $MV_{est2}$ (difference between MV and $MV_{est1}$ and difference between MV and $MV_{est2}$). Significance level was set at $P < 0.05$. IBM SPSS Statistics 20 software (LEAD Technologies, Inc., US) was used for statistical analyses.

## Results

*MV measured with MRI versus 3DUS*

There were no differences in MV measured from MRI as iopposed to 3DUS between the groups ($P = 0.371$). Thus, Bland-Altman analysis was performed with pooled data consisting of 18 participants (Fig. 2). The mean difference between the methods and linear regression analysis with a SEE are given in Table 1. The digitising error for MV (absolute difference between the same series of images digitised twice) was 2.1 ± 1.8% for 3DUS and 0.8 ± 1.7% for MRI.

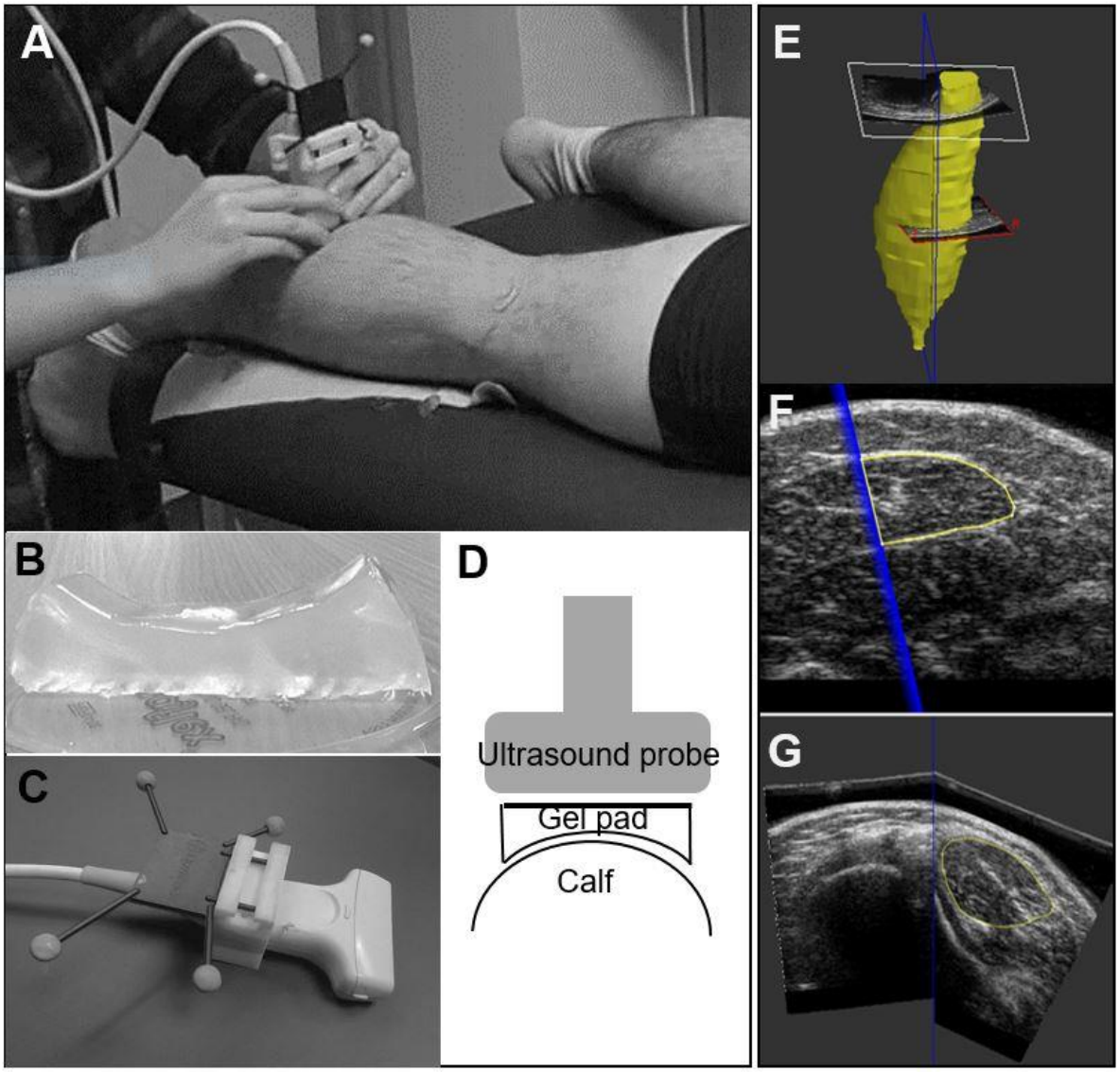


Figure 2. The assessment of muscle length with MRI and 3DUS methods. A sagittal view of the shank is shown to illustrate the muscle length assessment. 0 = anatomical medial gastrocnemius muscle origin, A = inferior margin of the medial condyle of the tibia, B = medial gastrocnemius muscle insertion at the musculotendinous junction, MG = medial gastrocnemius. Total muscle length from its anatomical origin to insertion is shown as a line from point 0 to point B on the MRI sagittal image. The lines connecting points A and B on both MRI and 3DUS sagittal images shows how muscle length was assessed for the 3DUS volume validity and comparison to Est1.

Table 1. Bland-Altman analysis for the comparison of measured muscle volume with MRI vs 3D ultrasound (MRI vs 3DUS), and between the measured and estimated techniques. The $R^2$ values with SEE from the linear regression analysis are given. MRI = muscle volume measured with MRI method, $Est_1$ = muscle volume estimated using $Est_1$ algorithm, $Est_2$ = muscle volume estimated using $Est_2$ algorithm.

| | Bland-Altman analysis | | | Regression analysis | |
|---|---|---|---|---|---|
| | Mean difference | UL of agreement | LL of agreement | $R^2$ | SEE% ($cm^3$) |
| MRI vs 3DUS | -0.13% (0.14 $cm^3$) | 3.75% (4.39 $cm^3$) | -4.02% (-4.10 $cm^3$) | 0.999 | 2.23 (1.924) |
| MRI vs $Est_1$ | 19.82% (25.14 $cm^3$) | 31.94% (62.03 $cm^3$) | 7.71% (-11.76 $cm^3$) | 0.990 | 4.73 (4.89) |
| MRI vs $Est_2$ | -3.11% (-2.72 $cm^3$) | 10.07% (14.09 $cm^3$) | -16.30% (-19.53 $cm^3$) | 0.988 | 6.55 (8.732) |

*Comparison of measured MV with $MV_{est1}$ and $MV_{est2}$*

The $ML_{MCT}$ and $ACSA_{max}$ measured with MRI as opposed to 3DUS were not different between groups. Pooled data from all groups (n = 18) showed no significant difference between MRI and 3DUS for mean $ML_{MCT}$ (MRI = 17.15 ± SD 2.89 cm and 3DUS = 17.16 ± 2.88 cm), mean $ACSA_{max}$ (MRI = 9.70 ± 4.44 $cm^2$ and 3DUS = 9.56 ± 4.07 $cm^2$), the mean $ACSA_{max}$ location (MRI = 56.09 ± 5.50% ranging from 41 to 67% and 3DUS = 56.41 ± 5.48% ranging from 48% to 71% along the $ML_{anatomical}$) and the mean $ACSA_{50\%}$ (MRI = 9.24 ± 2.81 $cm^2$ and 3DUS = 9.02 ± 4.01 $cm^2$).

The MV measured from 3DUS was compared to $MV_{est1}$ and $MV_{est2}$.. It was not possible to derive medial gastrocnemius $ML_{anatomical}$ from 3DUS measures in all participants, as the proximal MG origin was not clearly identifiable. Therefore, to compare our MV measures with $MV_{est2}$, $ML_{anatomical}$ was derived from MRI data. The one-way MANOVA revealed that the differences between the measured MV and $MV_{est1}$ and $MV_{est2}$ did not

differ between groups, thus the data were pooled for subsequent analyses (n=18). Bland-Altman analysis showed that the $MV_{est1}$ underestimated muscle volume by 19.82%. Despite a large systematic difference between the measured and estimated muscle volumes, the $MV_{est1}$ showed a strong linear relationship between the measured and estimated volumes ($R^2$ = 0.99)(Table 1). The $MV_{est2}$ slightly overestimated MV (mean difference = -3.11%). The mean differences and linear regression analysis with SEE are given in Table 1.

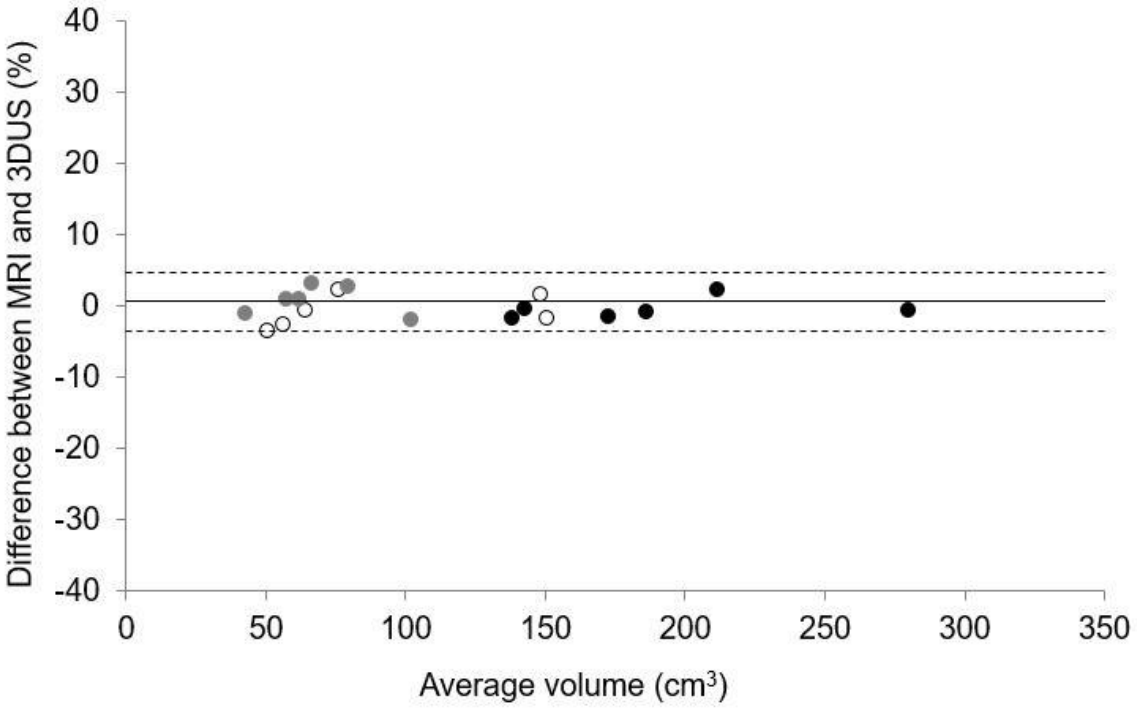


Figure 3. Bland-Altman plot for comparing muscle volume measured with MRI and 3DUS methods (n=18). The percentage difference between the methods is plotted against the mean volume measured with both methods. Mean differences and the upper and lower limits of agreement (bias) are shown. Black circles represent adults, white circles TD children and grey circles represent children with CP. MV = muscle volume.

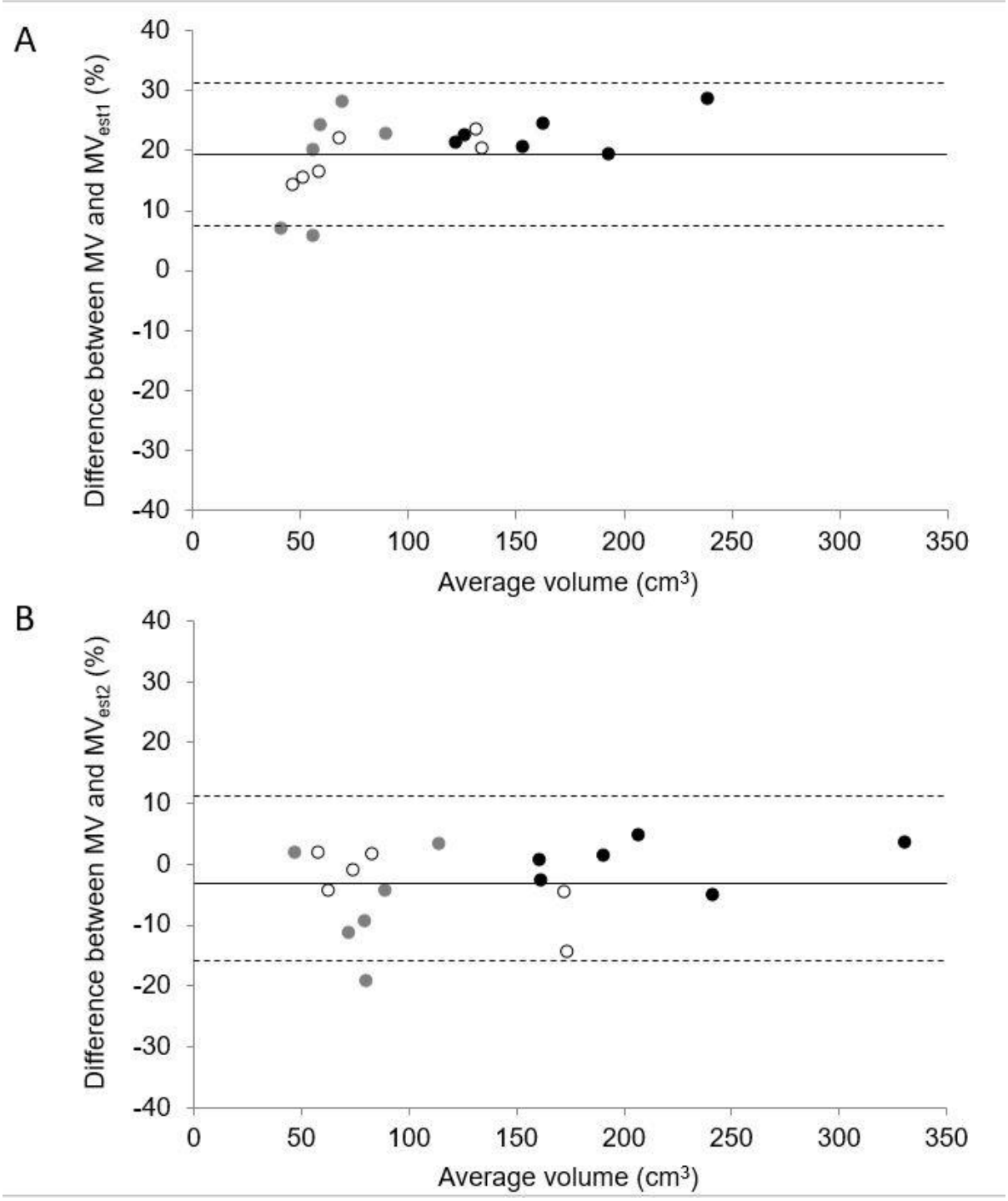


Figure 4. Bland-Altman plot for comparing measured and estimated muscle volumes (n=18). The percentage difference between the techniques is plotted against the mean volume measured with both $Est_1$ (4A) and $Est_2$ (4B) techniques. Mean differences and the upper and lower limits of agreement (bias) are shown. Black circles represent adults, white circles TD children and grey circles represent children with CP. MV = muscle volume, $Est_1$ = MV estimated with $Est_1$ method, $Est_2$ = MV estimated with $Est_2$ method.

## Discussion

The first purpose of the study was to examine the validity of MV measured with 3DUS compared to MRI in adults, TD children and children with CP. The results demonstrated that MV can be accurately measured with 3DUS in adults, TD children and children with CP. The second purpose was to assess the validity and generalisability of two estimation methods compared to MV using 3DUS. The results demonstrated good agreement with the $MV_{est2}$, however, there was a large systematic bias between measured volume and $MV_{est1}$, highlighting some methodological considerations.

The finding of good agreement between 3DUS and MRI in all groups is in line with a previous study by Barber et al. (2009) who found a similar 1.8 $cm^3$ (1.1%) difference in the level of agreement between the MRI and 3DUS, compared to 1.9 $cm^3$ shown in this study. Considering the ease and portability of the 3DUS method, in addition to shorter scanning procedures, 3DUS would seem to be a clear method of choice in clinical trials. However, careful consideration of methodological issues are needed for this method to be valid. For example, ultrasound imaging is limited by depth and field of view. This makes the image interpretation more difficult since anatomical landmarks such as bones and other surrounding muscles are not visible on the same scan. Additionally, the boundaries of muscle, particularly in very proximal locations, may not always be as visible compared to MRI scans. This notion is reflected in the values of digitising error for MRI (0.8%) and 3DUS (2.1%) in the present study. In this study, an average of two repeated digitised MV values were analysed and reported in order to minimise the digitising error. For novice users, it may be helpful to use a series of MRI images as a guide when learning and practicing the 3DUS scanning and digitising procedures. The importance of practicing the scanning procedure and having

anatomical knowledge of the muscle cannot be underestimated for accurate acquisition and digitising procedures from 3DUS data.

As expected, MV measured directly, were more accurate than volumes estimated with both algorithms. However, the $MV_{est2}$ differed by only -3.58% from the directly measured MV, with a SEE of 6.55% (8.73 $cm^3$). These results are similar to the results reported in the $MV_{est2}$ authors' original paper (Vanmechelena et al., 2018), where a SEE of 5.3% (9.7 $cm^3$) were found in young people with and without CP ($R^2$ = 0.998). Although the accuracy of the $MV_{est2}$ method was reasonably good, it is unclear if this is a less time-consuming alternative to a direct assessment of MV. In the present study, the location of the $ACSA_{max}$ was found to range between the 40-70% of $ML_{MCT}$. It is known that a small change in ACSA location can significantly change its value (Schless et al., 2017). Thus, using the $MV_{est2}$ method, one must firstly identify an individual's $ACSA_{max}$ before MV can be calculated. We also found it problematic to locate the proximal end of the medial gastrocnemius muscle in all participants in the present study using 3DUS imaging. A similar limitation has been described by Barber et al. (2009) who reported difficulty of visualising proximal insertion of the medial gastrocnemius in B-mode images. In order to compare measured MV to $MV_{est2,}$ total anatomical muscle length had to be derived from our MRI data. If it is not possible to calculate $ML_{anatomical}$ from the superficial aspect of the medial femoral condyle to the distal musculotendinous junction, the applicability and generalisability of this estimation method in a clinical setting may be limited.

The $MV_{est1}$ systematically underestimated muscle volume by ~19%. However, the linear regression showed a low SEE of 4.73.The authors of the $Est_1$ algorithm (Schless et al., 2017) found a mean absolute difference between measured and estimated

muscle volumes of 3.77 ± 2.9 $cm^3$ . In their study, the MV assessed with 3DUS was not compared to the MV measured with MRI, however, they reported a high repeatability for the ACSA assessment with 3DUS. One possible explanation for a large systematic underestimation of MV found in the present study could be a difference in the definition of the ACSA plane with respect to the 3D muscle reconstruction. In the present study, ACSA was measured as a cross-section, perpendicular to a line connecting the muscle's proximal and distal ends. However, in the method of Schless et al. (2017), the authors report using transverse images extracted from the 3D reconstructions, without defining the proximo-distal axial line of the 3D reconstructed muscle, relative to which the ACSA was perpendicular to. We may speculate that differences in the definition of the ACSA imaging plane could have introduced a systematic difference between the study by Schless et al. (2018) and the present study. This highlights that imaging plane and identification of anatomical landmarks are an important consideration when estimating MV. It is also noteworthy that for the $MV_{est1}$,the proximal origin of the medial gastrocnemius at the tibia, rather than the real anatomical origin at the femoral condyle, is being used. In the present study, this would have resulted in an underestimation of MV by 11.8 ± 2.8% compared to the MV measured for the whole anatomical length of the muscle.

The present study showed that estimations of MV using ML, ACSA and a form factor is a promising alternative to direct measures of MV. Difficulties in the measurement of $ML_{anatomical}$ and large ACSA with 2D ultrasound could be overcome with an extended field-of-view function in B-mode imaging. The extended field of view can provide an accurate measure of distance and ACSA when the probe angle in relation to the body surface is kept constant and follows a guided straight line (Noorkoiv et al., 2010). The validity and repeatability of this specific application, however, needs to be examined

further. Another consideration when estimating MV is that the method of estimation is based on the assumption that a muscle maintains a uniform shape along its length in response to exercise intervention or pathology. Several studies have challenged this assumption by demonstrating that the shape of the muscle changes with exercise in both trained and untrained individuals (Hedayatpour et al., 2012; Wakahara et al., 2013; Handsfield et al., 2017). The “shape/form factor” of different muscles in various populations and its response to various interventions, disability and disease should be explored further in future studies.

In conclusion, with sufficient expertise and improved image acquisition and scanning time by using the described gel pad, 3DUS is an accurate method for the assessment of medial gastrocnemius MV. The strong associations between MV, ACSA and ML suggest that the estimation methods have the potential to provide a simple, low cost and clinically feasible method for quantifying MV. Results of this paper, however, failed to replicate the method for estimating the $MV_{est1}$ showing a large systematic bias. The $MV_{est2}$ method does not appear to be less time consuming or less costly than the direct assessment of MV. Based on this study, the accurate assessment of muscle size, when using 3DUS, requires direct assessment instead of estimation. Example imaging data for practicing image acquisition and segmentation with 3DUS, in parallel with MRI, can be downloaded from https://figshare.com/s/04ba19d8a8f125eb2060.

**Acknowledgements**

The authors would like to thank the Dell Company, and i2 analytical for funding contributions.

**Conflicts of interest**

The authors have no conflicts of interest.

## References

Aagaard P, Suetta C, Caserotti P, Magnusson SP, Kjaer M. 2010. Role of the nervous system in sarcopenia and muscle atrophy with aging: strength training as a countermeasure. Scand J Med Sci Sports 20: 49-64.

Albracht K, Arampatzis A, Baltzopoulos B. 2008. Assessment of muscle volume and physiological cross-sectional area of the human triceps surae muscle in vivo. J Biomech 41(10): 2211-2218.

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.

Cerniglia LM, Delmonico MJ, Lindle R, Hurley BF, Rogers MA. 2007. Effects of acute supine rest on mid-thigh cross-sectional area as measured by computed tomography. Clin Physiol Funct Imaging 27(4): 249-253.

Delcker A, Walker F, Caress J, Hunt C, Tegeler C. 1999. In vitro measurement of muscle volume with 3-dimensional ultrasound. Eur J Ultrasound 9(2): 185-190.

Fry NR, Gough M, McNee AE, Shortland AP. 2007. Changes in the Volume and Length of the Medial Gastrocnemius After Surgical Recession in Children With Spastic Diplegic Cerebral Palsy. J Pediatr Ortop 27(7 ): 769-774.

Fukunaga, T. , Miyatani, M. , Tachi, M. , Kouzaki, M. , Kawakami, Y. and Kanehisa, H. 2017. Muscle volume is a major determinant of joint torque in humans. Acta Physiol Scand 172: 249–255.

Handsfield GG, Knaus KR, Fiorentino NM, Meyer CH, Hart JM, Blemker SS. 2017. Adding muscle where you need it: non-uniform hypertrophy patterns in elite sprinters. Scand J Med Sci Sports 27(10): 1050-1060.

Hedayatpour N, Falla D. 2009. Non-uniform muscle adaptations to eccentric exercise and the implications for training and sport. J Electromyogr Kinesiol 22(3): 329-333.

MacGillivray TJ, Ross E, Simpson HA, Greig CA. 2009. 3D freehand ultrasound for *in vivo* determination of human skeletal muscle volume. Ultrasound Med Biol 35(6): 928-35.

Malaiya R, McNee AE, Fry NR, Eve LC, Gough M, Shortland AP. 2007. The morphology of the medial gastrocnemius in typically developing children and children with spastic hemiplegic cerebral palsy." J Electromyogr and Kinesiol 17: 657-63.

Mersmann F, Bohm S, Schroll A, Arampatzis A. 2014. Validation of a simplified method for muscle volume assessment. J Biomech 47(6): 1348-52.

Narici MV, Roi GS, Landoni L, Minetti AE, Cerretelli P. 1989. Changes in force, cross-sectional area and neural activation during strength training and detraining of the human quadriceps. Eur J Occup Physiol 59(4): 310-319.

Noorkoiv M, Nosaka K, Blazevich AJ. 2010. Assessment of quadriceps muscle cross-sectional area by ultrasound extended-field-of-view imaging. Eur J Appl Physiol 109(4): 631-9.

O'Brien, T.D.,Reeves,N.D.,Baltzopoulos,V.,Jones,D.A.,Maganaris,C.N. 2009. Strong relationships exist between muscle volume, joint power and whole-body external mechanical power in adults and children. Exp Physiol 94: 731-738.

Prager RW, Rohling RN, Gee AH and Berman L. 1998. Rapid calibration for 3-D freehand ultrasound. Ultrasound Med Biol 24(6): 855-869.

Roig M, O'Brien K, Kirk G, Murray R, McKinnon P, Shadgan B, Reid WD. 2008. The effects of eccentric versus concentric resistance training on muscle strength and mass in healthy adults: a systematic review with meta-analyses. BJSM 43(8): 556-68.

Schless SH, Hanssen B, Cenni F, Bar-On L, Aertbelien E, Molenaers G, Desloovere K. 2017. Estimating medial gastrocnemius muscle volume in children with spastic cerebral palsy: a cross-sectional investigation. DMCN 60(1): 81-87.

Treece GM, Gee AH, Prager RW, Cash CJC and Berman L. 2003. High definition freehand 3D ultrasound. Ultrasound Med Biol 29(4): 529-546.

Treece GM, Prager RW, Gee AH and Berman L. 2002. Correction of probe pressure artifacts in freehand 3D ultrasound. Med Image Anal 6(3): 199-214.

Vanmechelena IM, Shortlanda AP, Noble JJ. 2018. Lower limb muscle volume estimation from maximum cross-sectional area. Clin Biomech 51: 40-44.

Wakahara T, Fukutani A, Kawakami Y, Yanai T. 2013. Nonuniform muscle hypertrophy: its relation to muscle activation in training session. Med Sci Sports Exerc 45(11): 2158-65.

Weide G, van der Zwaard S, Huijing PA, Jaspers RT, Harlaar J. 2017. 3D Ultrasound Imaging: Fast and Cost-effective Morphometry of Musculoskeletal Tissue. J Vis Exp 129.

Weller R, Pfau T, Ferrari M, Griffith R, Bradford T, Wilson A. 2007. The determination of muscle volume with a freehand 3D ultrasonography system. Ultrasound Med Biol 33(3): 402-407.