QUANTIFICATION OF THIGH MUSCLE VOLUME IN CHILDREN AND ADOLESCENTS USING MAGNETIC RESONANCE IMAGING

Short Title: Accuracy of volumetric calculation methods in youth

Owen William Tomlinson [1, 2], Alan Robert Barker [1], Jonathan Fulford [3], Paul Wilson [1], Patrick John Oades [2], Craig Anthony Williams [1, 2]

Affiliations

1. Children's Health and Exercise Research Centre, Sport and Health Sciences, University of Exeter, Heavitree Road, Exeter, EX1 2LU, United Kingdom.
2. Royal Devon and Exeter NHS Foundation Trust Hospital, Barrack Road, Exeter, EX2 5DW, United Kingdom.
3. Institute of Biomedical and Clinical Science, University of Exeter Medical School, University of Exeter, Heavitree Road, Exeter, EX1 2LU, United Kingdom.

Corresponding author:

Professor Craig Williams

Children's Health and Exercise Research Centre,

Sport and Health Sciences,

University of Exeter,

Exeter, EX1 2LU

UNITED KINGDOM

Tel: +44 (0)1392 724890

Email: C.A.Williams@exeter.ac.uk

ABSTRACT

Estimating muscle volume (MV) using variable numbers of cross-sectional area (CSA) slices obtained from magnetic resonance imaging (MRI) introduces error that is known in adults, but not in children and adolescents, whereby body sizes differ due to growth and maturation. Therefore, 15 children and adolescents (11 males, 14.8 ± 2.1 years) underwent MRI scans of the right thigh using a 1.5 T scanner to establish this error. A criterion MV was determined by tracing around and summing all CSAs, with MV subsequently estimated using every second, third, fourth and fifth CSA slice. Bland-Altman plots identified mean bias and limits of agreement (LoA) between methods. Error rates between 1.0 – 10.4% were seen between criterion and estimated MV. Additional analyses identified an impact of formulae selection, with a cylindrical formula preferred to a truncated cone. To counter high error between criterion and estimated MV due to discrepancies in the number of CSA slices analysed, length-matched criterion volumes were established, with reduced error rates (0.5 – 2.0%) being produced as a result. CSA at 50% thigh length also predicted MV, producing high error (13.8 – 39.6%). Pearson's correlation coefficients determined relationships between error and measures of body size/composition, with all body size/composition measures being correlated (r = -0.78 – 0.86, $p < 0.05$) with the error between criterion and estimated MV. To conclude, MV can be accurately estimated using fewer CSA slices. However, the associated error must be considered when calculating MV in children and adolescents, as body size biases estimates.

KEYWORDS: adolescence, cross-sectional area, musculoskeletal, limits of agreement, respiratory disease.

INTRODUCTION

Accurate quantification and interpretation of muscle size is important in physiological studies, such as those measuring hypertrophy following training (Tracy et al., 1999), muscle atrophy following immobilisation (Wall et al., 2014) or aging (Ogawa, Yasuda, & Abe, 2012), and examining the consequences of chronic disease (Godi et al., 2016).

To quantify muscle volume (MV), magnetic resonance imaging (MRI) is considered the preferred technique due to use of non-ionising radiation, while producing high resolution images (Narici, Landoni, & Minetti, 1992), and consists of the measurement and summation of multiple sequential cross-sectional areas (CSA) (Barnouin et al., 2015; Tracy et al., 2003). As this is time consuming, studies have sought to identify the measurement error associated with increasing the distance between measured CSAs with the objective of reducing the number of CSAs required and the time taken for analysis (Barnouin, et al., 2015; Tracy, et al., 2003; Walton, Roberts, & Whitehouse, 1997). However, as the number of CSA slices decreases, the associated error with estimated MV increases. For example, Tracy, et al. (2003) reports the limits of agreement (LoA) increasing from ±0.7% to ±6.4% of total MV when 11 mm and 91 mm gaps between CSA slices are used.

Further considerations are a) the direction in which CSA slices are sequentially summed to estimate MV, and b) the choice of geometric model used to calculate MV. Previous studies have estimated MV from the knee, working towards the hip (i.e. distal to proximal [D-P]) (Nordez et al., 2009; Tracy, et al., 2003), a process that may under-estimate thigh volume. This is likely due to the actual shape of the thigh, whereby a wide proximal

circumference and narrow distal circumference are present, producing a broadly conical model. Such a D-P method assumes a smaller distal slice accurately reflects the size of larger proximal slices, whereas this is not actually the case – thus resulting in underestimation of MV. Furthermore, previous research in adults has assessed differing calculation methods, such as the cylindrical and truncated cone formulae (Barnouin, et al., 2015). However, no study has systematically evaluated whether the direction of measurement (i.e. D-P, or a proximal-to-distal [P-D] direction) has a bearing on final MV estimates, nor whether choice of geometric model influences calculation of MV in youth.

To our knowledge, studies examining the errors associated with determining MV have only been undertaken in adults. Therefore, the measurement strategies applied may not be suitable for groups involving children and adolescents. Compared to adults, children and adolescents have a different body geometry (Feber and Krásničanová, 2012) and the process of maturation (timing and tempo of maturity) leads to children and/or adolescents of equal chronological age, but different body size (Mirwald, Baxter-Jones, Bailey, & Beunen, 2002) and MV (Pitcher et al., 2012). These factors are likely to influence the error when determining MV using MRI, and warrant further investigation.

The recognition of measurement errors are vital in clinical populations. For example, Duchenne muscular dystrophy, where progressive MV decline can arise and has previously been quantified using MRI (Godi, et al., 2016). Similarly, nutritional complications in cystic fibrosis (CF) lead to considerable variations in body-size (Culhane, George, Pearo, & Spoede, 2013), and recent debate has queried whether a qualitative or quantitative muscular defect is predominantly responsible for impaired

oxidative metabolism (Hulzebos, Jeneson, van der Ent, & Takken, 2017; Rodriguez-Miguelez, Erickson, McCully, & Harris, 2017). Within CF, previous studies have utilised only muscle CSA from a single slice (e.g. at 50% of limb length) to reflect muscle size (Moser, Tirakitsoontorn, Nussbaum, Newcomb, & Cooper, 2000). However, whilst single site CSA is a poor surrogate for total MV in healthy adults (Morse, Degens, & Jones, 2007), this has yet to be examined in clinical and non-clinical groups of children and adolescents.

The primary aim of this study was to identify the error associated with estimating MV from MRI using a differing number of CSA slices in two groups of children and adolescents; healthy controls (CON), and a group with CF. Secondary aims included: a) identifying the difference in estimated MV when employing a P-D or D-P approach to analysing CSA slices; b) identify the relationship between body size and the error in quantifying MV; and c) identify the utility of mid-thigh CSA to predict MV.

METHODS

Study Population

Fifteen children and adolescents (8 CF [2 female, 6 male], 7 healthy controls [CON; 2 female, 5 male], 14.8 ± 2.1 years) volunteered for the study, with descriptive characteristics presented in Table 1. Individuals with CF were recruited if they satisfied inclusion criteria (diagnosis based on clinical features and where possible, genotyping; lung function [forced expiratory volume in one-second, FEV_1] considered stable and within 10% of best in preceding six months; and no increase in symptoms or weight loss in preceding two weeks). Subsequently, age- and sex-matched controls were recruited

based upon characteristics of those in the CF group. All participants were included if they understood the study protocol and presented no contraindications to being within a scanner environment (e.g. metallic implants, claustrophobia).

All participants were recruited from a hospital CF clinic, local schools and sports clubs. Ethics approval was obtained from NHS Regional Ethics Committee (14/SW/0061), and participants and parents/guardians provided written informed assent and consent respectively.

Anthropometric Measures

Stature, seated stature and leg length (i.e. stature – seated stature) were obtained using wall-mounted and seated stadiometers (Holtain, Crymych, Wales) to the nearest 0.1 cm. Body mass (BM) was measured to the nearest 0.01 kg (Seca, Birmingham, UK). Skinfold callipers and published equations (Slaughter et al., 1988) were used to estimate body fat percentage, which was used to determine fat mass and fat-free mass (FFM).

Quantification of Volume

MV of the right thigh was determined using a 1.5 T superconducting whole-body scanner (Gyroscan Intera, Philips, the Netherlands), utilising a T1 weighted image sequence to obtain a series of transverse slices covering the whole upper leg with optimal fat/muscle signal contrast. Participants lay in the prone position within the scanner, with the hips and upper legs extended and secured to avoid unnecessary movement, but not to cause compression of muscle tissue.

Slices were acquired with 5 mm thickness and 0.5 mm slice gap, similar to previous research (Barnouin, et al., 2015; Nordez, et al., 2009). CSA was determined using Philips software, by manually tracing around the muscle within each slice. The CSA value for each individual slice, apart from the first and last was multiplied by 5.5 mm (5 mm slice thickness + 0.5 mm slice gap) to produce individual slice volumes. Slices at the distal and proximal ends were multiplied by 5.25 mm to reflect the absence of the 0.25 mm contribution from the adjacent slice gap. All individual volumes were then summed over all slices to calculate a criterion measure of MV.

CSA analyses were undertaken by two investigators, with a within-investigator coefficient of variation (CV) < 1.5% and a between-investigator CV of 1.2%. The number of CSA slices required to cover the length of the thigh ranged from 40 to 56.

Selection of Formulae

Multiple formulae are available with which to quantify MV, including the cylindrical and truncated cone formula. The cylindrical formula assumes a cylindrical shape for the muscle and is presented by the following equation:

$$MV = \sum_{n} h \times CSA_i \qquad [1]$$

whereby: MV = muscle volume, *n* = number of slices used, h = slice thickness, CSA = cross-sectional area of slice (Nordez, et al., 2009) and has been used previously (Lund et al., 2002; Walton, et al., 1997). The truncated cone formula (Ross, Rissanen, Pedwell, Clifford, & Shragge, 1996);

$$MV = \sum_{n-1} {}^{h}/_{3} \times (CSA_i + CSA_{i+1} + \sqrt{CSA_i \times CSA_{i+1}}) \qquad [2]$$

has also been utilised in previous studies, although has been shown to produce a higher level of error compared to the cylindrical method in adults (Barnouin, et al., 2015; Nordez, et al., 2009). However, as this error between methods has yet to be established in a paediatric population, a preliminary assessment was undertaken to compare error between the two formulae to determine which formula to use in further analyses (see Preliminary Analyses in Results).

Slicing Intervals

In accordance with previous research (Tracy, et al., 2003), estimated MV was calculated by increasing the interval between CSA slices using every second (MV2), third (MV3), fourth (MV4) and fifth (MV5) slice. For MV2 – MV5, each slice CSA, apart from for the first and last slice (which lacked an adjacent 0.25 mm as previously described), was multiplied by 11, 16.5, 22 or 27.5 mm respectively, prior to summing over all slices to produce an estimate for MV.

These slicing intervals could potentially create instances whereby an estimated MV could be compared to the true criterion MV consisting of a different number of CSA slices. For example, the shortest thigh length assessed in this study consisted of 40 slices, and therefore use of the MV5 strategy would utilise slices 1, 6, 11, 16, 21, 26, 31 and 36 to create an estimate of MV. Subsequently, a comparison between a true criterion value of 40 slices and estimated volume of 36 slices could possibly introduce further error. To determine the value of the error associated with comparing the true criterion MV against estimated MV of differing lengths, another preliminary assessment was undertaken to quantify such bias (see Preliminary Analyses in Results).

In such instances when the number of CSA slices covering the upper leg did not exactly fit the slice sampling frequency, a reduced number of slices were examined in line with previous research (Tracey et al. 2003), to create a 'length matched criterion' (LMC) MV. Using the previous example of 40 CSA slices with the MV5 method, the estimated MV using 36 slices would be compared against a LMC MV, also of 36 slices (i.e. missing the remaining four CSA slices from either the proximal or distal end, dependent of direction of analysis). Such an approach ensured that any observed bias was due to slicing intervals and analysis direction, and not differences in assessed thigh lengths.

Direction of Analysis

The above mentioned procedures were completed for both the D-P and P-D directions, and therefore for the described example, in the opposite direction, slices 40, 35, 30, 25, 20, 15, 10, and 5 would be used to estimate MV5. This estimated MV would then be compared against a new LMC MV of 36 slices, using each slice from 40-5.

Estimation of MV from mid-thigh CSA

In addition, the CSA of the slice obtained that lay nearest to 50% of the length of the measured femur (i.e. midpoint between femoral head and medial epicondyle) was assessed to determine the predictive ability of a single slice in estimating true criterion MV. The CSA of the whole thigh (muscle, subcutaneous fat, intramuscular fat infiltration and connective tissue; CSA_T) and muscle-only CSA (no subcutaneous fat, intramuscular fat infiltration and where possible and visible in scans, no connective tissue; CSA_M) were

recorded. Bone was excluded from all CSA slices, as were blood vessels when visible on scans.

Statistical Analyses

Pearson's correlation coefficients and paired samples *t*-tests were conducted to identify relationships and mean differences, respectively, between LMC MVs, and each of the MV estimates (i.e. MV2 to MV5). Bland-Altman plots (Bland and Altman, 1986) were produced to identify the mean bias between measures of MV (MV and MV2, MV3, MV4, MV5 for both D-P and P-D directions), and associated 95% LoA. Furthermore, Bland-Altman analyses were conducted between predicted MV from both CSA_T and CSA_M. The MV predicted from CSA_T and CSA_M was obtained using simple bi-linear regression. For each Bland-Altman plot, Pearson's correlations between means (*x*-axis) and differences (*y*-axis) were conducted to identify whether the size and direction of the error is associated with estimated MV itself.

To identify if the over- or under-estimation of MV is associated with anthropometric variables, the absolute difference between each MV estimate and LMC MVs was calculated and correlated against chronological age, as well as body size parameters (stature, leg length, body mass, FFM, fat mass) using Pearson's correlation coefficients. To account for multiple comparisons, a Benjamini-Hochberg correction was utilised. All statistical analyses were performed using SPSS v.23 (IBM Corp., Armonk, NY, USA), with statistical significance taken at an alpha value of 0.05. Effect sizes (*ES*) were described for mean comparisons (small = 0.2, medium = 0.5, large = 0.8) and correlation coefficients (small = 0.1, medium = 0.3, large = 0.5) (Cohen, 1992).

RESULTS

Participants

Participants with CF within this study presented with differing genotypes (ΔF508/ ΔF508 = 4; ΔF508/Unknown = 1; ΔF508/711+1G->T = 1; ΔF508/E585X = 1). No differences were observed between individuals with, and without, CF for any anthropometric, MV or CSA variables ($p > 0.05$, Table 1). Furthermore, FEV_1 was not different between groups (CF, 104 ± 11 vs. CON, 98 ± 22 $\%_{Predicted}$, $p = 0.51$, $ES = 0.35$) indicating preserved function in CF and equivalence between groups. Therefore, all variables are pooled (i.e. $n = 15$) for subsequent analyses.

Preliminary Analyses

Preliminary analysis of the truncated cone formula identified a mean bias of 11 – 55 cm^3, and error rate of 0.4 – 2.1% (Supplementary Figure 1). Whilst this error was comparable to that of the cylindrical formula (Figure 1), correlation coefficients of a greater magnitude were reported between the differences between, and means of, estimated MV and LMC MV, relative to the cylindrical formula, indicating bias in the error (Supplementary Table 1).

Comparisons between the true criterion MV and respective estimates identified a mean bias of 27 – 275 cm^3, and associated error rate of 1.0 – 10.4%, with notable differences dependent on direction of analysis (Supplementary Figure 2). Furthermore, variances are observed in both the direction and magnitude of correlations between the differences between, and means of, estimated and true criterion MV (Supplementary Table 1).

263

264 *Main Results*

265 As preliminary analyses identified the use of the cylindrical formula, using LMC MV
266 values was most appropriate and held lowest potential for error, these were carried
267 forward into the main analyses. Mean values (with associated LoA) for the difference
268 from each LMC MV for all estimation methods, using the cylindrical formula, and
269 directions are displayed in Figure 1. The mean LMC MV was significantly greater than
270 each estimated volume using the D-P slicing direction, and significantly lower than each
271 estimated volume for the P-D direction (all $p < 0.001$, $ES = 0.03 – 0.12$). All estimated
272 MV (P-D and D-P) variables were significantly, and positively correlated with their LMC
273 MV (all $r = 1.0$, all $p < 0.001$). Furthermore, the mean bias and LoA associated with each
274 estimation method increased as the interval between slices increases (see Figure 1 using
275 Bland-Altman plots). When using the D-P direction, a significant and positive correlation
276 was evident between the mean difference between LMC and estimated MV, and the
277 respective means for MV3, MV4 and MV5, but not MV2 (Supplementary Table 1). For
278 P-D, significant and negative correlations are observed between the means and
279 differences of LMC and estimated MV for all slicing strategies (Supplementary Table 1).

280

281 Both CSA_M and CSA_T were significant predictors of MV (Figure 2). When each
282 predictive equation was used to estimate MV, mean bias was equal to zero for both CSA
283 parameters, with LoA for CSA_M (384 cm^3, 13.8%) being smaller than CSA_T (1099 cm^3,
284 39.6%; Figure 1). The correlations between the mean and difference of the true criterion
285 and estimated MV for CSA_M ($r = 0.12$, $p = 0.66$) and CSA_T ($r = 0.43$, $p = 0.13$) were
286 positive, but not statistically significant.

Significant correlations were found between age and body size values, and both the absolute and percentage difference between LMC and estimated MV (Table 2) for both the D-P (r = 0.13 – 0.86, r = -0.04 – -0.51) and P-D (r = -0.20 – -0.78, r = -0.29 – 0.75) directions. Differences from true criterion MV estimated using CSA_M were associated with leg length, and estimates using CSA_T were associated with stature, leg length, body fat percentage, FFM and fat mass. The only slicing strategy to not hold any significant correlations (in either absolute or percentage terms) was MV2 in the D-P direction. Leg length had moderate correlations (r > 0.3) with all slicing estimate differences apart from the absolute error of MV2.

DISCUSSION

This study confirms, in children and adolescents, that as the interval between slices increases (and therefore the number of CSA slices decreases), the mean bias and 95% LoA associated with the error also increases. In addition, it has been shown that the direction of slicing affects the magnitude of the mean bias and associated LoA, although for both P-D and D-P directions, as MV increases, as does the error associated with their respective over- and under-estimation. Furthermore, our results have established the error associated with using a single CSA slice to predict MV, and the relationships between measures of body size and differences between LMC MV and estimated MV. These findings are novel for children and adolescents, adding to previous work conducted in adults (Barnouin, et al., 2015; Tracy, et al., 2003), and therefore have implications for the accurate determination of MV in individuals that present different morphology to adults, including those with chronic disease.

Within our investigation, we performed a series of preliminary analyses, establishing for the first time in a paediatric population the error associated with using the truncated cone formula and subsequent correlations between the differences and means between estimated and LMC MV, as shown in Supplementary File 1. Whilst the error was similar to the cylindrical formula, in contrast to work in adults (Barnouin, et al., 2015), correlations were of a greater magnitude, indicating a systematic bias in the estimation methods that could alter final estimates of MV. Furthermore, the bias associated between the true criterion value and that of each MV estimate was established, with error rates up to 10.4% using MV5 being identified (Supplementary Figure 2). Consequently, these preliminary findings resulted in the remaining analyses utilising the cylindrical formula, and adoption of a LMC method, as per previous research (Tracy, et al., 2003), in order to retain as low an error and systematic bias as possible, thus isolating the effect of slicing strategy and analysis direction within final results.

In the current study, all MV estimates were significant correlated with their LMC MV – a finding that would initially indicate a high level of agreement. However, as correlations do not indicate systematic biases within an agreement, use of Bland-Altman analyses have been utilised, and identified that as the number of slices decreased (i.e. from MV2 to MV5), the error associated with estimating MV increased – in agreement with research in adults (Barnouin, et al., 2015; Nordez, et al., 2009; Tracy, et al., 2003; Walton, et al., 1997). Within this finding, the greatest bias was evident at MV5 (27.5 mm gap; LoA = ±2.0% of MV), a finding that is similar to Tracy, et al. (2003), who utilised a slice gap of 31 mm (Tracy, et al., 2003), and resulted in a LoA of ±1.7% of MV. Previous studies,

however, have only been conducted in adults, and have only sought to identify the MV of the quadriceps femoris (QF) group, whereas the present study used the MV of the whole thigh in a group of children and adolescents. Studies investigating structure and functional relationships should consider utilising whole-thigh MV (as the current study has done), as research has identified equal recruitment of both quadriceps and hamstrings during cycling exercise (Richardson, Frank, & Haseler, 1998); however, this will be dependent on individual research questions being investigated.

Whilst the error established in this study could be considered small (a maximal error rate of ±2.0%, using LMC, as per Figure 1), the acceptability of such error is dependent on the research question being addressed. In cross-sectional studies assessing differences in MV between groups with disease (Mathur, Takai, Macintyre, & Reid, 2008), or age differences in healthy participants (Maden-Wilkinson, McPhee, Rittweger, Jones, & Degens, 2014; Tolfrey et al., 2006), such error would be consistently applied across both groups and therefore such estimation methods could be acceptably utilised. However, interventional studies investigating temporal changes in MV (i.e. atrophy, hypertrophy) may be required to detect changes that may fall inside the margins of error established in the current study. For example, a bed-rest study in adults from Belavy et al. (2009) identified reductions in MV in individual thigh muscles ranging from 9 cm^3 (7.3%; biceps femoris) to 34 cm^3 (12.3%; semimembranosus) following 56 days of immobilisation; such changes in MV may not be detected when a larger inter-slice distance is used. Furthermore, an additional challenge is posed in this decision making, whereby the results of the preliminary analyses must also be considered. For example, when considering the difference between the true criterion and estimated value (Supplementary Figure 2),

researchers and clinicians must decide whether to wholly exclude 'remaining' slices if they should exist; to measure them separately; or to accept the bias relative to the true criterion, especially if only using slicing strategy such as MV2 which produced a maximal error of 2.8%. Therefore, the acceptable slice interval (and choice of methodology) will be dependent on the level of precision needed in any outcome variables.

A new outcome in this study is that the measurement error is dependent upon the direction of measurement (i.e. D-P vs. P-D), which has not previously been assessed in children and adolescents. In previous work (Nordez, et al., 2009; Tracy, et al., 2003), measures from the knee towards the hip (D-P), have under-estimated MV. The findings of the current study identified a reduced LoA observed in the P-D direction compared to D-P. This difference was greatest between estimates using MV5, where a difference in LoA of 15 cm^3 (0.6%) was reported. This difference between directional approaches is likely due to the use of a simple cylindrical approach to estimating MV, by assuming each inter-slice volume is appropriately represented by the initial CSA. Due to the broadly conical shape of the thigh (i.e. wide proximal end, and narrow distal end), this approach will result in under-estimation using a D-P direction, and an over-estimation in a P-D direction.

Whilst the mean bias and LoA do show a difference for the slicing directions, given the magnitude of previously described MV changes following interventions (Belavy, et al., 2009), the direction of measurement is unlikely to have a clinically meaningful impact upon final MV estimates. However, as shown in Figure 1, the significant correlation between the means and differences of each LMC and estimated MV measure using the P-

383 D direction suggests that use of this direction may be biased. The only MV measure that
384 did not identify a significant correlation is MV2 using the D-P direction, indicating this
385 option may be the most suitable MV estimate in the current study.
386
387 A further noteworthy finding within this study is the association between body size and
388 the absolute and percentage differences between LMC and estimated MV (Table 2).
389 These findings suggest that a child's age and body size can impact the final MV estimates,
390 and have implications when heterogeneous groups of children are being assessed (e.g.
391 those with variances in age, stature and mass). Results show that when the error is
392 presented as a percentage of LMC MV (to minimise further bias by muscle size), the error
393 associated with estimates using MV2 and MV3 in the D-P direction do not provide
394 significant correlations with body size and may therefore be suitable for future use as they
395 are not biased by the range of different body sizes within children and adolescents (unlike
396 MV4 and MV5). This difference between estimation methods is likely due to the
397 resolution in estimating MV (high resolution in MV2, low resolution in MV5) which
398 introduces the technical errors described within the study.
399
400 Of note, leg length held medium correlations ($r > 0.3$) with all MV errors when expressed
401 as a percentage, further confirming our concerns regarding biasing of estimates due to
402 body size. This association was shown to be significant when greater inter-slice distances
403 are used (i.e. MV5), and therefore this may result in an upper-limit to the value of the
404 inter-slice distance used when estimating MV. This is of further concern when studies
405 utilise a fixed number of CSA slices to calculate MV (Mathur, et al., 2008; Nordez, et al.,
406 2009) as this can result in a varying inter-slice distance for each participant dependent on

the size of the limb being investigated. This is of concern in studies involving children and adolescents, where body size is heterogeneous, as evidenced by the use of between 40 and 56 CSA slices per participant to calculate the criterion MV in the current study. Therefore, when the relationship between error and body size, and the possible evidence for an upper limit between slices is considered, use of a fixed number of slices may not provide a uniform amount of bias across participants in studies calculating MV. Therefore, this approach cannot be recommended for use without an appropriate comparison of the respective methodologies (i.e. fixed inter-slice distance vs. fixed number of slices).

The number of studies using MRI to undertake MV calculations in disease groups is limited, with Duchenne muscular dystrophy (Godi, et al., 2016), chronic obstructive pulmonary disorder (Mathur, et al., 2008) and CF (Moser, et al., 2000) utilising this methodology. In individuals with CF, we are aware of only one previous study that has utilised MRI to identify muscle size, which used mid-thigh CSA to infer reduced exercise capacity (Moser, et al., 2000). However, the use of a single CSA slice has been shown to be a poor predictor (R^2 = 0.79, SEE = 27%) of MV in adults (Morse, et al., 2007). The current study agreed with Morse, et al. (2007) in identifying a significant relationship between CSA and MV (Figure 2). Whilst the shared variance between these variables (R^2 = 0.53, 0.94) would initially indicate a predictive ability, a large SEE was also identified, with nearly 40% error being reported as the LoA for MV predicted from CSA_T and over 13% for CSA_M. These errors are over twenty-times, and seven-times, the size of the largest LoA for MV5 reported in the present study, respectively. Therefore, whilst the use of a single CSA slice is a time-efficient method in comparison to summation of multiple

CSA slices, the magnitude of error observed suggests estimation from a single CSA slice is not a valid method for determining MV and should be discouraged.

Within the present study, a lack of difference between groups (CF and CON) was observed for parameters of both body, and muscle, size. When considered alongside the relatively preserved FEV_1 of 104 ± 11 $\%_{Predicted}$, it could be considered that the CF cohort were in fact a 'healthy disease' group, which in turn accounts for the lack of differences between groups, which has been observed for CSA parameter previously (Moser, et al., 2000). This lack of differences may limit the external validity of our findings to other individuals with CF, although such homogeneity across the whole group has enhanced the power behind the current analyses for a combined child/adolescent group.

In summary, when quantifying MV in children and adolescents using CSA slices obtained from MRI scans, this study has identified: a) an increased error when the intervals between slices is increased: b) an influence of the direction in which MV is estimated: c) the poor predictive ability of a single CSA slice to estimate MV, and d) when a slice interval of MV3 and above is used, the resultant differences are related to body size. These findings lead to a practical recommendation that use of MV2 in the D-P direction may be suitable for estimating MV in children and adolescents as: a) it halves the time required for analysis whilst, b) the resultant error does not hold a relationship with body size parameters, nor is systematically biased by the mean of the criterion and estimated MV itself.

FUNDING

The facility requirements for this study were financially supported by Sport and Health Sciences at the University of Exeter and the Royal Devon & Exeter CF Research Charitable Fund. Jonathan Fulford's salary was supported via an NIHR grant to the University of Exeter (CRF/2016/10027). No further external funding is reported. There are no conflicts of interest to report.

ACKNOWLEDGEMENTS

The authors would thank all participants and parents who volunteered for this study. Furthermore, appreciation is extended to staff and coaches from the Royal Devon & Exeter NHS Foundation Trust Cystic Fibrosis Team, Exmouth Community College and CrossFit Exe for assistance with recruitment.

CONFLICT OF INTEREST

The authors have no conflicts of interest.

REFERENCES

Barnouin, Y., Butler-Browne, G., Moraux, A., Reversat, D., Leroux, G., Béhin, A., . . . Hogrel, J.-Y. (2015). Comparison of different methods to estimate the volume of the quadriceps femoris muscles using MRI. *Journal of Medical Imaging and Health Informatics, 5*(6), pp. 1201-1207. doi:10.1166/jmihi.2015.1506

Belavy, D. L., Miokovic, T., Armbrecht, G., Richardson, C. A., Rittweger, J., & Felsenberg, D. (2009). Differential atrophy of the lower-limb musculature during prolonged bed-rest. *European Journal of Applied Physiology, 107*(4), pp. 489-499. doi:10.1007/s00421-009-1136-0

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), pp. 307-310. doi:10.1016/S0140-6736(86)90837-8

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), pp. 155-159. doi:10.1037/0033-2909.112.1.155

Culhane, S., George, C., Pearo, B., & Spoede, E. (2013). Malnutrition in cystic fibrosis: a review. *Nutrition in Clinical Practice, 28*(6), pp. 676-683. doi:10.1177/0884533613507086

Feber, J., & Krásničanová, H. (2012). Measures of body surface area in children. In V. Preedy (Ed.), *Handbook of Anthropometry* (pp. 1249-1256). New York NY, USA: Springer.

Godi, C., Ambrosi, A., Nicastro, F., Previtali, S. C., Santarosa, C., Napolitano, S., . . . Politi, L. S. (2016). Longitudinal MRI quantification of muscle degeneration in Duchenne muscular dystrophy. *Annals of Clinical and Translational Neurology, 3*(8), pp. 607-622. doi:10.1002/acn3.319

Hulzebos, H. J., Jeneson, J. A., van der Ent, C. K., & Takken, T. (2017). CrossTalk opposing view: Skeletal muscle oxidative capacity is not altered in cystic fibrosis patients. *Journal of Physiology, 595*(5), pp. 1427-1428. doi:10.1113/JP272505

Lund, H., Christensen, L., Savnik, A., Boesen, J., Danneskiold-Samsoe, B., & Bliddal, H. (2002). Volume estimation of extensor muscles of the lower leg based on MR imaging. *European Radiology, 12*(12), pp. 2982-2987. doi:10.1007/s00330-002-1334-1

Maden-Wilkinson, T. M., McPhee, J. S., Rittweger, J., Jones, D. A., & Degens, H. (2014). Thigh muscle volume in relation to age, sex and femur volume. *Age, 36*(1), pp. 383-393. doi:10.1007/s11357-013-9571-6

Mathur, S., Takai, K. P., Macintyre, D. L., & Reid, D. (2008). Estimation of thigh muscle mass with magnetic resonance imaging in older adults and people with chronic obstructive pulmonary disease. *Physical Therapy, 88*(2), pp. 219-230. doi:10.2522/ptj.20070052

Mirwald, R. L., Baxter-Jones, A. D. G., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements *Medicine and Science in Sports and Exercise, 34*(4), pp. 689-694. doi:10.1097/00005768-200204000-00020

Morse, C. I., Degens, H., & Jones, D. A. (2007). The validity of estimating quadriceps volume from single MRI cross-sections in young men. *European Journal of Applied Physiology, 100*(3), pp. 267-274. doi:10.1007/s00421-007-0429-4

Moser, C., Tirakitsoontorn, P., Nussbaum, E., Newcomb, R., & Cooper, D. M. (2000). Muscle size and cardiorespiratory response to exercise in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine, 162*(5), pp. 1823-1827. doi:10.1164/ajrccm.162.5.2003057

Narici, M. V., Landoni, L., & Minetti, A. E. (1992). Assessment of human knee extensor muscles stress from in vivo physiological cross-sectional area and strength measurements. *European Journal of Applied Physiology and Occupational Physiology, 65*(5), pp. 438-444.

Nordez, A., Jolivet, E., Sudhoff, I., Bonneau, D., de Guise, J. A., & Skalli, W. (2009). Comparison of methods to assess quadriceps muscle volume using magnetic resonance imaging. *Journal of Magnetic Resonance Imaging, 30*(5), pp. 1116-1123. doi:10.1002/jmri.21867

Ogawa, M., Yasuda, T., & Abe, T. (2012). Component characteristics of thigh muscle volume in young and older healthy men. *Clinical Physiology and Functional Imaging, 32*(2), pp. 89-93. doi:10.1111/j.1475-097X.2011.01057.x

Pitcher, C. A., Elliott, C. M., Williams, S. A., Licari, M. K., Kuenzel, A., Shipman, P. J., . . . Reid, S. L. (2012). Childhood muscle morphology and strength: alterations

over six months of growth. *Muscle & Nerve, 46*(3), pp. 360-366. doi:10.1002/mus.23326

Richardson, R. S., Frank, L. R., & Haseler, L. J. (1998). Dynamic knee-extensor and cycle exercise: functional MRI of muscular activity. *Int J Sports Med, 19*(3), pp. 182-187. doi:10.1055/s-2007-971901

Rodriguez-Miguelez, P., Erickson, M. L., McCully, K. K., & Harris, R. A. (2017). CrossTalk proposal: Skeletal muscle oxidative capacity is altered in patients with cystic fibrosis. *Journal of Physiology, 595*(5), pp. 1423-1425. doi:10.1113/JP272486

Ross, R., Rissanen, J., Pedwell, H., Clifford, J., & Shragge, P. (1996). Influence of diet and exercise on skeletal muscle and visceral adipose tissue in men. *Journal of Applied Physiology, 81*(6), pp. 2445-2455. doi:10.1152/jappl.1996.81.6.2445

Slaughter, M. H., Lohman, T. G., Boileau, R. A., Horswill, C. A., Stillman, R. J., Vanloan, M. D., & Bemben, D. A. (1988). Skinfold equations for estimation of body fatness in children and youth. *Human Biology, 60*(5), pp. 709-723.

Tolfrey, K., Barker, A., Thom, J. M., Morse, C. I., Narici, M. V., & Batterham, A. M. (2006). Scaling of maximal oxygen uptake by lower leg muscle volume in boys and men. *Journal of Applied Physiology, 100*(6), pp. 1851-1856. doi:10.1152/japplphysiol.01213.2005

Tracy, B. L., Ivey, F. M., Hurlbut, D., Martel, G. F., Lemmer, J. T., Siegel, E. L., . . . Hurley, B. F. (1999). Muscle quality. II. Effects of strength training in 65- to 75-yr-old men and women. *Journal of Applied Physiology, 86*(1), pp. 195-201.

Tracy, B. L., Ivey, F. M., Jeffrey Metter, E., Fleg, J. L., Siegel, E. L., & Hurley, B. F. (2003). A more efficient magnetic resonance imaging-based strategy for measuring quadriceps muscle volume. *Medicine and Science in Sports and Exercise, 35*(3), pp. 425-433. doi:10.1249/01.MSS.0000053722.53302.D6

Wall, B. T., Dirks, M. L., Snijders, T., Senden, J. M., Dolmans, J., & van Loon, L. J. (2014). Substantial skeletal muscle loss occurs during only 5 days of disuse. *Acta Physiologica, 210*(3), pp. 600-611. doi:10.1111/apha.12190

Walton, J. M., Roberts, N., & Whitehouse, G. H. (1997). Measurement of the quadriceps femoris muscle using magnetic resonance and ultrasound imaging. *British Journal of Sports Medicine, 31*(1), pp. 59-64. doi:10.1136/bjsm.31.1.59

573 **Table 1.** Descriptive characteristics of anthropometric and MRI derived variables, and
574 differences between CF and CON groups.

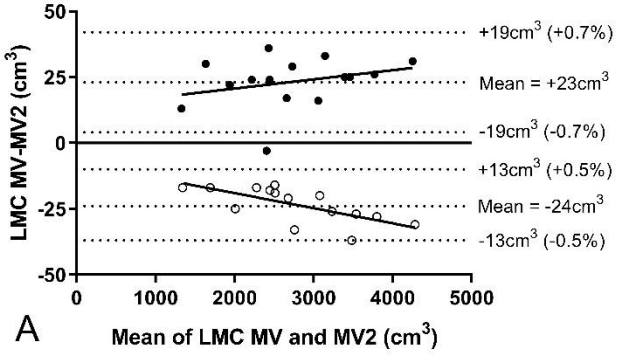
Variable	Combined (n = 15)	CF (n = 8)	CON (n = 7)	p Value	*ES*
Age (years)	14.8 (2.1)	15.1 (2.1)	14.4 (2.2)	0.57	0.31
Stature (m)	1.62 (0.11)	1.63 (0.11)	1.62 (0.11)	0.84	0.09
Sitting stature (m)	0.85 (0.56)	0.85 (0.06)	0.84 (0.06)	0.63	0.33
Leg length (m)	0.78 (0.52)	0.77 (0.05)	0.78 (0.06)	0.89	0.18
Body mass (kg)	57.22 (15.45)	61.68 (18.02)	52.13 (11.03)	0.25	0.63
Body fat (%)	17.8 (6.1)	18.6 (6.3)	16.7 (6.1)	0.61	0.31
FFM (kg)	46.47 (11.45)	49.30 (12.47)	43.30 (10.14)	0.33	0.52
Fat mass (kg)	10.75 (6.51)	12.43 (8.12)	8.83 (3.74)	0.30	0.56
True criterion MV (cm^3)	2778 (801)	2823 (763)	2726 (901)	0.83	0.12
CSA_M (cm^2)	59.67 (15.57)	62.75 (13.99)	56.15 (17.61)	0.43	0.42
CSA_T (cm^2)	93.17 (24.16)	98.86 (25.64)	86.66 (22.41)	0.35	0.50
CSA_M (% of CSA_T)	64.7 (10.1)	64.6 (11.4)	64.7 (9.5)	0.98	0.01

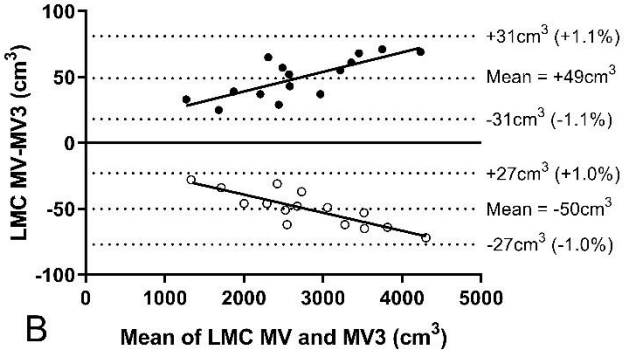
575 Data are presented as mean ± standard deviation. CF, cystic fibrosis; CON, control; FFM,
576 fat-free mass; MV, muscle volume; CSA_M, muscle cross-sectional of area at mid-thigh;
577 CSA_T, cross-sectional area of muscle and subcutaneous fat at mid-thigh; *ES*, effect size.
578
579
580
581
582
583
584
585

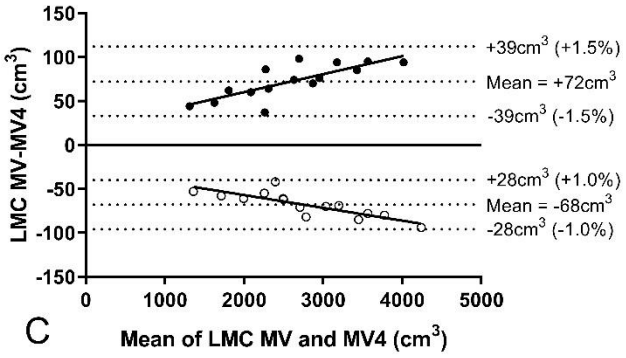
Table 2. Pearson's correlation coefficients between differences of each estimate of muscle volume (MV*n*) and respective length matched criterion muscle volume, and body size variables.

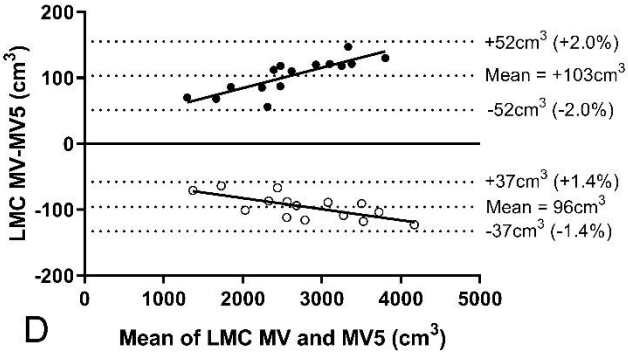
	Age (years)	Stature (m)	Leg Length (m)	Body Mass (kg)	Body Fat (%)	FFM (kg)	Fat Mass (kg)
Distal to Proximal							
$MV2_a$	r = 0.33, p = 0.36	r = 0.19, p = 0.63	r = 0.14, p = 0.72	r = 0.13, p = 0.75	r = 0.16, p = 0.68	r = 0.13, p = 0.73	r = 0.62, p = 0.87
$MV2_{\%}$	r = -0.25, p = 0.49	r = -0.46, p = 0.17	r = -0.47, p = 0.17	r = -0.39, p = 0.26	r = 0.22, p = 0.55	r = -0.48, p = 0.16	r = -0.07, p = 0.85
$MV3_a$	**r = 0.73, p = 0.02**	**r = 0.64, p = 0.04**	r = 0.55, p = 0.09	**r = 0.65, p = 0.04**	r = 0.03, p = 0.93	**r = 0.70, p = 0.02**	r = 0.31, p = 0.38
$MV3_{\%}$	r = -0.08, p = 0.84	r = -0.37, p = 0.30	r = -0.36, p = 0.30	r = -0.15, p = 0.71	r = 0.17, p = 0.67	r = -0.28, p = 0.43	r = -0.14, p = 0.72
$MV4_a$	**r = 0.86, p < 0.01**	**r = 0.70, p = 0.02**	r = 0.56, p = 0.09	**r = 0.71, p = 0.02**	r = 0.31, p = 0.38	**r = 0.70, p = 0.02**	r = 0.47, p = 0.17
$MV4_{\%}$	r = -0.04, p = 0.92	r = -0.40, p = 0.25	r = -0.43, p = 0.21	r = -0.15, p = 0.70	r = 0.56, p = 0.08	r = -0.39, p = 0.26	r = -0.32, p = 0.37
$MV5_a$	**r = 0.80, p < 0.01**	**r = 0.78, p = 0.01**	r = 0.61, p = 0.05	**r = 0.83, p < 0.01**	r = 0.32, p = 0.38	**r = 0.82, p < 0.01**	r = 0.53, p = 0.11
$MV5_{\%}$	r = -0.09, p = 0.83	r = -0.40, p = 0.25	**r = -0.51, p = 0.02**	r = -0.07, p = 0.85	r = 0.56, p = 0.09	r = -0.30, p = 0.39	r = -0.37, p = 0.30
Proximal to Distal							
$MV2_a$	**r = -0.64, p = 0.04**	**r = -0.67, p = 0.03**	r = -0.49, p = 0.16	**r = -0.74, p = 0.01**	r = -0.43, p = 0.21	**r = -0.69, p = 0.02**	r = -0.55, p = 0.11
$MV2_{\%}$	r = 0.32, p = 0.37	r = 0.46, p = 0.17	r = 0.57, p = 0.08	r = 0.18, p = 0.64	r = -0.55, p = 0.10	r = 0.41, p = 0.24	r = -0.29, p = 0.42
$MV3_a$	**r = -0.70, p = 0.02**	**r = -0.76, p = 0.01**	**r = -0.72, p = 0.02**	**r = -0.65, p = 0.04**	r = 0.07, p = 0.85	**r = -0.76, p = 0.01**	r = -0.20, p = 0.61
$MV3_{\%}$	r = 0.27, p = 0.45	r = 0.42, p = 0.21	r = 0.33, p = 0.36	r = 0.31, p = 0.39	r = -0.09, p = 0.83	r = 0.40, p = 0.25	r = 0.03, p = 0.93
$MV4_a$	**r = -0.78, p = 0.01**	**r = -0.76, p = 0.01**	r = -0.57, p = 0.08	**r = -0.74, p = 0.02**	r = -0.24, p = 0.52	**r = -0.76, p = 0.01**	r = -0.41, p = 0.24
$MV4_{\%}$	r = 0.47, p = 0.17	**r = 0.68, p = 0.02**	**r = 0.75, p = 0.01**	r = 0.44, p = 0.20	r = -0.34, p = 0.35	**r = 0.63, p = 0.04**	r = -0.06, p = 0.87
$MV5_a$	**r = -0.69, p = 0.02**	**r = -0.71, p = 0.02**	**r = -0.63, p = 0.04**	**r = -0.65, p = 0.04**	r = -0.30, p = 0.40	**r = -0.64, p = 0.04**	r = -0.42, p = 0.23
$MV5_{\%}$	r = 0.42, p = 0.19	**r = 0.65, p = 0.04**	r = 0.61, p = 0.05	r = 0.47, p = 0.17	r = -0.35, p = 0.34	**r = 0.66, p = 0.03**	r = -0.06, p = 0.10
Mid-point Cross Sectional Areas							
CSA_{Ma}	r = 0.21, p = 0.57	r = 0.27, p = 0.45	**r = 0.73, p = 0.02**	r = -0.05, p = 0.88	r = -0.16, p = 0.68	r = 0.03, p = 0.93	r = 0.23, p = 0.54
$CSA_{M\%}$	r = 0.28, p = 0.44	r = 0.40, p = 0.25	r = 0.55, p = 0.09	r = 0.09, p = 0.82	r = -0.09, p = 0.82	r = 0.17, p = 0.67	r = -0.07, p = 0.84
CSA_{Ta}	r = 0.25, p = 0.49	r = 0.56, p = 0.09	r = 0.50, p = 0.15	r = 0.10, p = 0.82	**r = -0.71, p = 0.02**	r = 0.44, p = 0.20	**r = 0.76, p = 0.01**
$CSA_{T\%}$	r = 0.32, p = 0.37	**r = 0.67, p = 0.03**	**r = 0.79, p = 0.01**	r = 0.23, p = 0.52	r = -0.57, p = 0.08	r = 0.53, p = 0.10	r = -0.39, p = 0.27

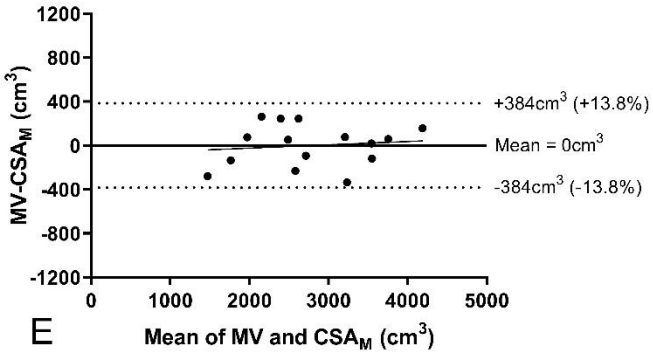
MV*n*, estimated muscle volume using every n^{th} CSA slice; CSA_M, muscle only cross-sectional area of mid-thigh; CSA_T, whole thigh cross-sectional area of mid-thigh; FFM, fat-free mass. Significant results ($p < 0.05$) following Benjamini-Hochberg correction are highlighted in bold. Subscript 'a' and '%' indicate whether error is expressed as an absolute value or percentage.

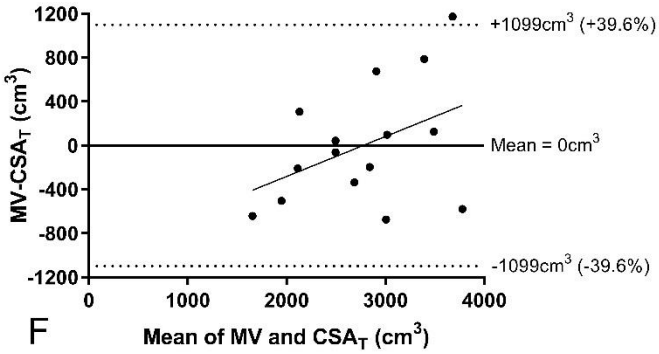












591

592

593

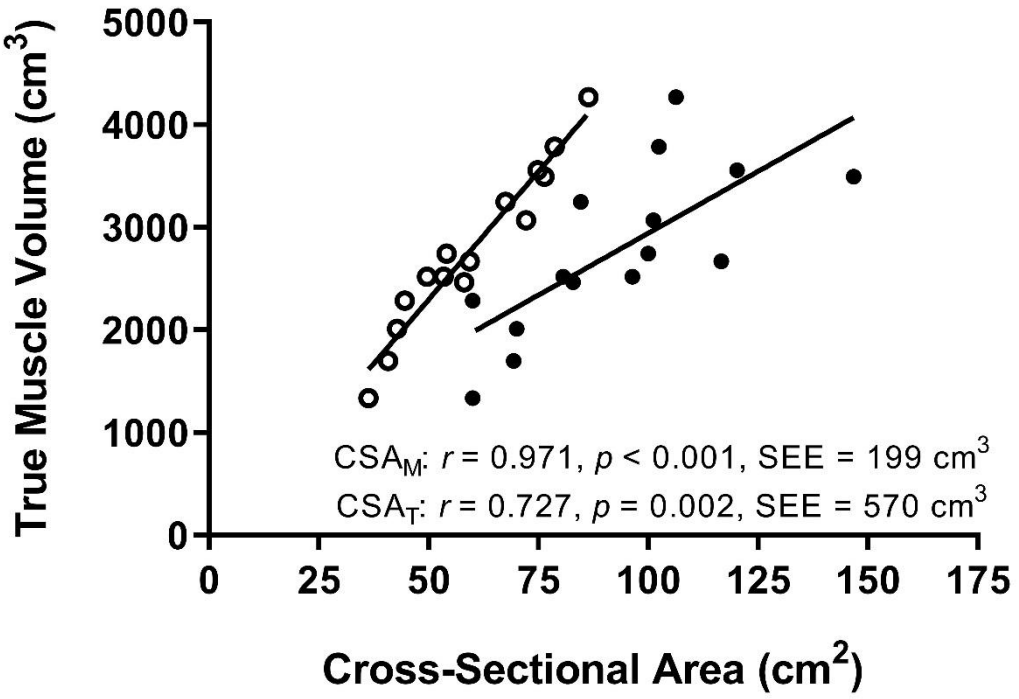
594

595

596

597

Figure 1. Bland-Altman plots identifying relationships between the differences between (*y*–axis) and mean of (*x*–axis) estimated uscle volume (MV*n*) and length matched criterion muscle volume (LMC MV). Plots display use of different slicing strategies (A = MV2, B = MV3, C = MV4, D = MV5) and directions (D-P = black circles, P-D = white circles). Predicted MV from CSA_M and CSA_T are in plots E and F respectively. All plots show: mean bias (central dashed horizontal line); 95% limits of agreement limits (±2 standard deviations; upper and lower dashed horizontal lines) presented as absolute values (cm^3) and as a percentage of LMC MV; correlation between means and differences (solid diagonal lines) for each MV estimate. MV*n*, estimated muscle volume using every n^{th} CSA slice; CSA_M, muscle only cross-sectional area of mid-thigh; CSA_T, whole thigh cross-sectional area of mid-thigh.



622

623 **Figure 2.** The relationship between true criterion muscle volume and CSA_M (white
624 circles) and CSA_T (black circles). CSA_M, muscle only cross-sectional area of mid-thigh;
625 CSA_T, whole thigh cross-sectional area of mid-thigh; *r*, Pearson's correlation coefficient;
626 *p*, significance value; SEE = standard error of the estimate.

627

628

629

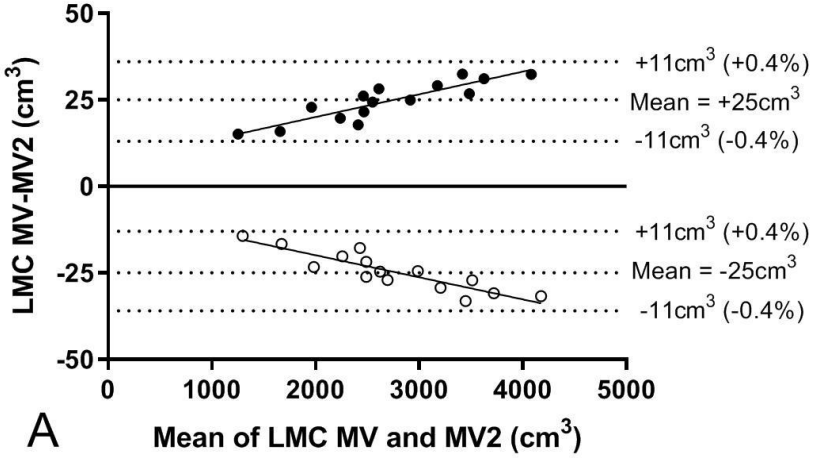
630

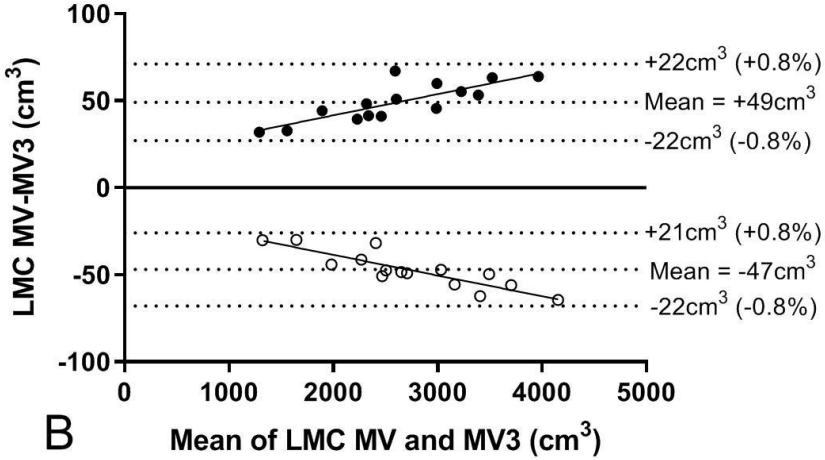
631

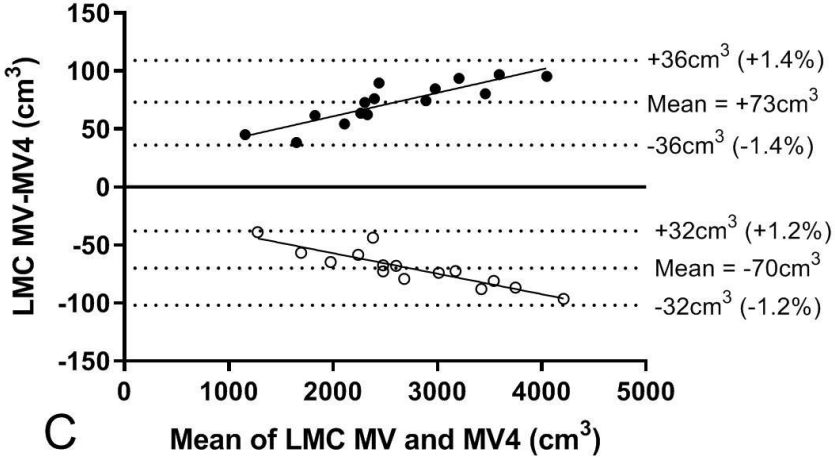
632

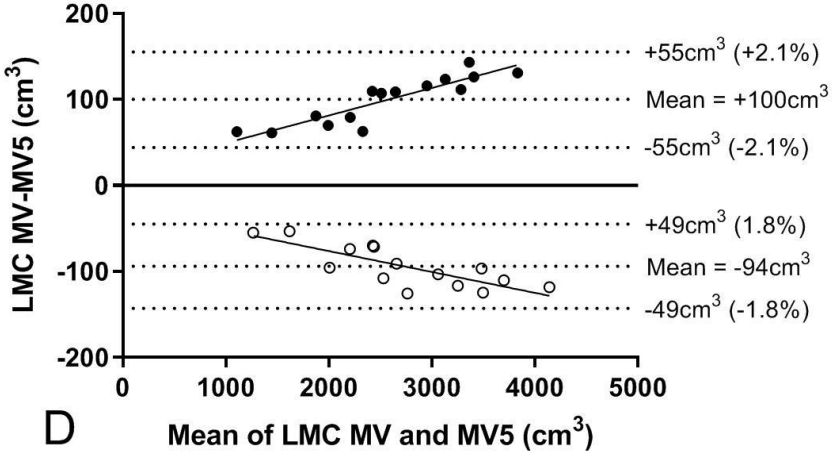
633

634


LMC MV-MV2 (cm3)
+11cm3 (+0.4%)
Mean = +25cm3
-11cm3 (-0.4%)
+11cm3 (+0.4%)
Mean = -25cm3
-11cm3 (-0.4%)
Mean of LMC MV and MV2 (cm3)
A


LMC MV-MV3 (cm3)
+22cm3 (+0.8%)
Mean = +49cm3
-22cm3 (-0.8%)
+21cm3 (+0.8%)
Mean = -47cm3
-22cm3 (-0.8%)
Mean of LMC MV and MV3 (cm3)
B


LMC MV-MV4 (cm3)
+36cm3 (+1.4%)
Mean = +73cm3
-36cm3 (-1.4%)
+32cm3 (+1.2%)
Mean = -70cm3
-32cm3 (-1.2%)
Mean of LMC MV and MV4 (cm3)
C


LMC MV-MV5 (cm3)
+55cm3 (+2.1%)
Mean = +100cm3
-55cm3 (-2.1%)
+49cm3 (1.8%)
Mean = -94cm3
-49cm3 (-1.8%)
Mean of LMC MV and MV5 (cm3)
D

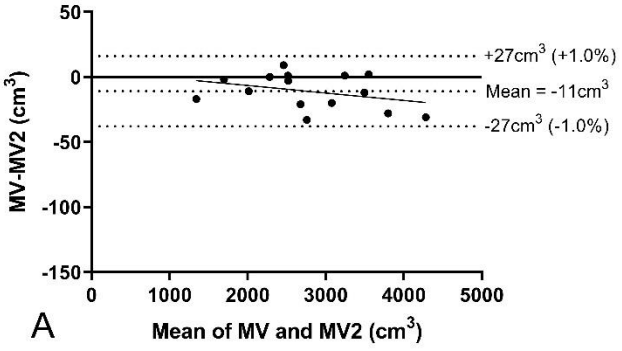
635

Supplementary Figure 1. Bland-Altman plots identifying relationships between the differences between (*y*–axis) and mean of (*x*–axis) estimated muscle volume (MV*n*) and respective length matched criterion muscle volume (LMC MV) using the truncated cone formula to estimate MV. Plots display use of different slicing strategies (A = MV2, B = MV3, C = MV4, D = MV5) and directions (D-P = black circles, P-D = white circles).

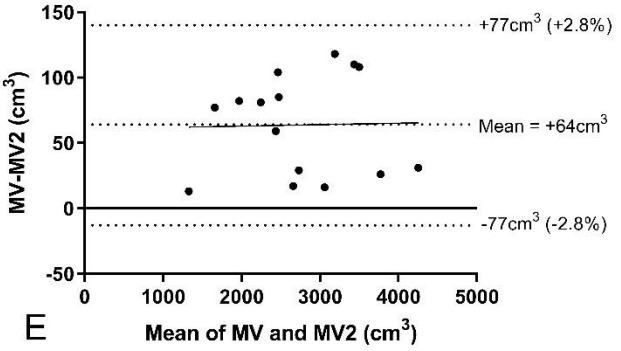
Values above zero (i.e. positive bias) indicate the estimated volume is less than the criterion value and therefore a consistent under-estimation is present. Values below zero (i.e. negative bias) indicate the estimated volume is greater than the criterion value and therefore a consistent over-estimation is present.

All plots show: mean bias (central dashed horizontal line); 95% limits of agreement limits (±2 standard deviations; upper and lower dashed horizontal lines) presented as absolute values (cm^3) and as a percentage of LMC MV.

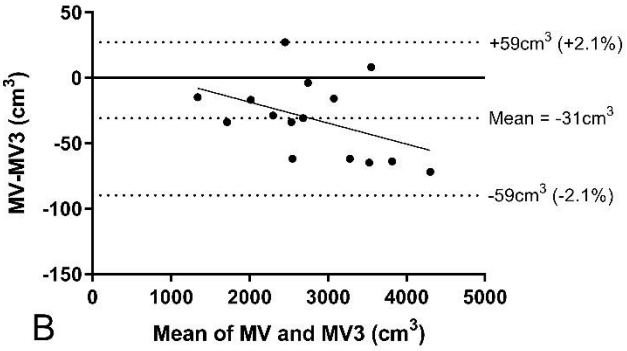
Correlation between means and differences (solid diagonal lines) are presented for each MV estimate using Pearson's coefficients, with respective *r* and *p* values reported in Supplementary Table 1. MV*n*, estimated muscle volume using every n^{th} CSA slice.


+27cm3 (+1.0%)
Mean = -11cm3
-27cm3 (-1.0%)
MV-MV2 (cm3)
Mean of MV and MV2 (cm3)

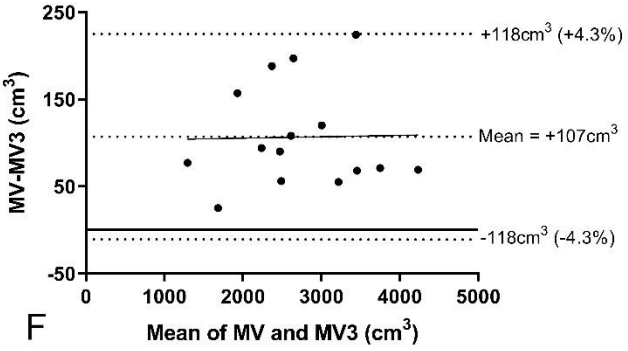
A


+77cm3 (+2.8%)
Mean = +64cm3
-77cm3 (-2.8%)
MV-MV2 (cm3)
Mean of MV and MV2 (cm3)

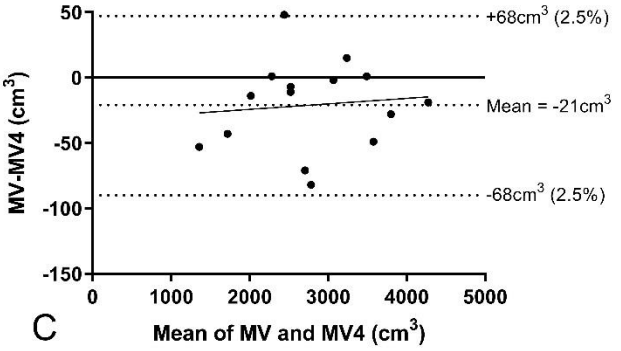
E


+59cm3 (+2.1%)
Mean = -31cm3
-59cm3 (-2.1%)
MV-MV3 (cm3)
Mean of MV and MV3 (cm3)

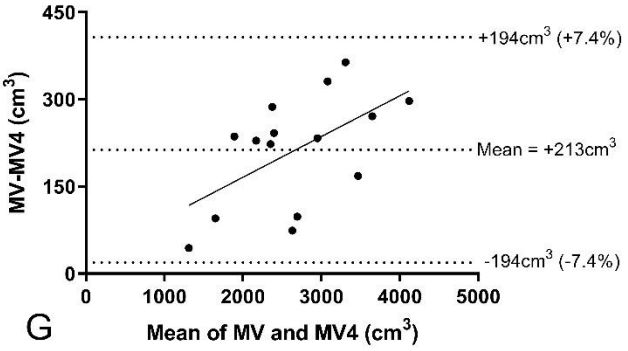
B


+118cm3 (+4.3%)
Mean = +107cm3
-118cm3 (-4.3%)
MV-MV3 (cm3)
Mean of MV and MV3 (cm3)

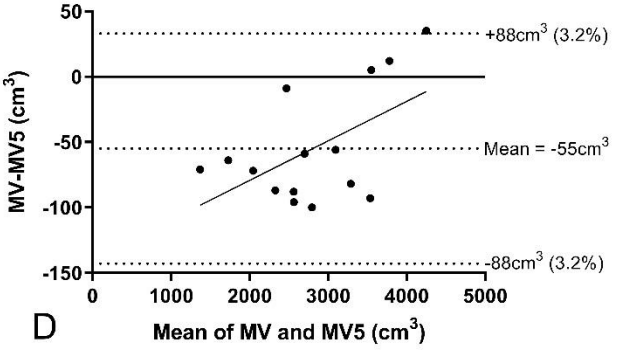
F


+68cm3 (2.5%)
Mean = -21cm3
-68cm3 (2.5%)
MV-MV4 (cm3)
Mean of MV and MV4 (cm3)

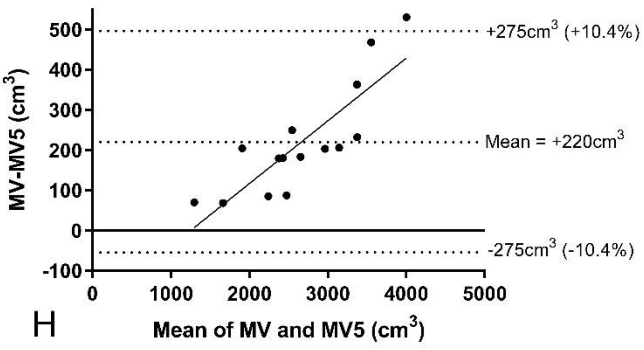
C


+194cm3 (+7.4%)
Mean = +213cm3
-194cm3 (-7.4%)
MV-MV4 (cm3)
Mean of MV and MV4 (cm3)

G


+88cm3 (3.2%)
Mean = -55cm3
-88cm3 (3.2%)
MV-MV5 (cm3)
Mean of MV and MV5 (cm3)

D


+275cm3 (+10.4%)
Mean = +220cm3
-275cm3 (-10.4%)
MV-MV5 (cm3)
Mean of MV and MV5 (cm3)

H

Supplementary Figure 2. Bland-Altman plots identifying relationships between the differences between (*y*–axis) and mean of (*x*–axis) estimated muscle volume (MV*n*) and true criterion MV. Plots display use of different directions (A-D = proximal to distal, E-H = distal to proximal). All analyses have been calculated using a cylindrical equation for determination of MV.

Values above zero (i.e. positive bias) indicate the estimated volume is less than the criterion value and therefore a consistent under-estimation is present. Values below zero (i.e. negative bias) indicate the estimated volume is greater than the criterion value and therefore a consistent over-estimation is present.

All plots show mean bias (central dashed horizontal line) and 95% limits of agreement limits (±2 standard deviations; upper and lower dashed horizontal lines). Values are presented in absolute terms (cm^3) and as a percentage of true criterion MV.

Correlation between means and differences (solid diagonal lines) are presented for each MV estimate using Pearson's coefficients, with respective *r* and *p* values reported in Supplementary Table 1. MV*n*, estimated muscle volume using every n^{th} CSA slice.

Supplementary Table 1. Pearson's correlation coefficients for the relationship between the differences between, and means of, estimated and criterion muscle volumes using differing calculation methods and directions of calculation.

	D-P	P-D
Cylindrical Formula		
MV2	$r = 0.290$, $p = 0.295$	$r = -0.691$, $p = 0.004$
MV3	$r = 0.767$, $p = 0.001$	$r = -0.829$, $p < 0.001$
MV4	$r = 0.783$, $p = 0.001$	$r = -0.827$, $p < 0.001$
MV5	$r = 0.823$, $p < 0.001$	$r = -0.697$, $p = 0.004$
Truncated Cone Formula		
MV2	$r = 0.897$, $p < 0.001$	$r = -0.897$, $p < 0.001$
MV3	$r = 0.809$, $p < 0.001$	$r = -0.875$, $p < 0.001$
MV4	$r = 0.865$, $p < 0.001$	$r = -0.884$, $p < 0.001$
MV5	$r = 0.897$, $p < 0.001$	$r = -0.789$, $p < 0.001$
True Criterion Volume		
MV2	$r = 0.023$, $p = 0.935$	$r = -0.344$, $p = 0.209$
MV3	$r = 0.022$, $p = 0.939$	$r = -0.439$, $p = 0.101$
MV4	$r = 0.559$, $p = 0.030$	$r = -0.098$, $p = 0.727$
MV5	$r = 0.840$, $p < 0.001$	$r = -0.543$, $p = 0.036$

MV*n*, estimated muscle volume using every n^{th} CSA slice. D-P, distal to proximal direction; P-D, proximal to distal direction. Cylindrical and truncated cone analyses were made using length matched criterion values.