Estimation of changes in volume of individual lower-limb muscles using magnetic resonance imaging (during bed-rest)

D L Belavý[1]**, T Miokovic**[1]**, J Rittweger**[2,3]** and D Felsenberg**[1]

[1] Charité Universitätsmedizin Berlin, Center for Muscle and Bone Research, Hindenburgdamm 30, 12200 Berlin, Germany
[2] Institute for Biomedical Research into Human Movement and Health, Manchester Metropolitan University, Manchester, M1 5GD, UK
[3] Institute of Aerospace Medicine, German Aerospace Center, 51147 Cologne, Germany

E-mail: belavy@gmail.com

Abstract

Muscle size in the lower limb is commonly assessed in neuromuscular research as it correlates with muscle function and some approaches have been assessed for their ability to provide valid estimates of muscle volume. Work to date has not examined the ability of different measurement approaches (such as cross-sectional area (CSA) measures on magnetic resonance (MR) imaging) to accurately track changes in muscle volume as a result of an intervention, such as exercise, injury or disuse. Here we assess whether (a) the percentage change in muscle CSA in 17 lower-limb muscles during 56 days bed-rest, as assessed by five different algorithms, lies within 0.5% of the muscle volume change and (b) the variability of the outcome measure is comparable to that of muscle volume. We find that an approach selecting the MR image with the highest muscle CSA and then a series of CSA measures, the number of which depended upon the muscle considered, immediately distal and proximal, provided an acceptable estimate of the muscle volume change. In the vastii, peroneal, sartorius and anterior tibial muscle groups, accurate results can be attained by increasing the spacing between CSA measures, thus reducing the total number of MR images and hence the measurement time. In the two heads of biceps femoris, semimembranosus and gracilis, it is not possible to reduce the number of CSA measures and the entire muscle volume must be evaluated. Using these approaches one can reduce the number of CSA measures required to estimate changes in muscle volume by ∼60%. These findings help to attain more efficient means to track muscle volume changes in interventional studies.

Keywords: magnetic resonance imaging, skeletal muscle, cross-sectional area, exercise, measurement

S Online supplementary data available from stacks.iop.org/PM/32/35/mmedia

1. Introduction

Muscle volume is commonly used as an indicator of muscle size. Muscle size can also be taken to mean its cross-sectional area (CSA), either at a specific anatomical point or from a series of points along the length of a muscle. The size of skeletal musculature is frequently measured as part of human research in a number of fields as it correlates with a variety of functional parameters (Blazevich *et al* 2009, Fukunaga *et al* 2001, Trappe *et al* 2001). A number of factors determine the functional characteristics (e.g. force, power or stiffness) of individual muscles. Aside from neuromuscular control, fibre-type characteristics and muscle fibre pennation angle, one of the factors is the amount of muscle bulk present (Narici *et al* 1989). Furthermore, muscle size measurements are often used as a surrogate outcome parameter for understanding muscle function where measuring individual muscle force output *in vivo* and/or electromyography are not possible or appropriate. Thus, the measurement of muscle size, often via magnetic resonance imaging (MRI), is widely used in neuromuscular research.

One of the difficulties encountered when using this parameter is that the measurement of the 'size' of a muscle (i.e. its entire volume) can be very time consuming and technically difficult. Typically, a series of MR images are taken from a body region and then an operator must manually trace around the border of the muscle(s) of interest, calculate the area and repeat the process for subsequent MR images of a given muscle. In our experience, the measurement of muscle volume of the lower-limb musculature in one leg of a single subject based upon 65 images requires approximately 10 h of operator time. This can vary depending upon the number of MR images present. Given that investigations of the musculature commonly involve multiple subjects measured at multiple time-points, it is easy to see how such measurements can become very time consuming. Whilst some semi-automated MR image analysis packages exist (e.g. Brem *et al* 2009) which can potentially reduce operator time for the measurement of certain muscles, their use commonly relies upon stark contrast differences in the MR images to enable the software to delineate muscle borders. Depending on the quality of images obtained and the particular muscle being imaged, clear delineation of muscle borders is not always possible. Hence, the underlying problem of operator time required to measure multiple images and subsequently estimate muscle size remains. One method which would assist in the reduction of operator time would be to decrease the number of images analysed. However, this raises the question of how correctly sub-sets of the image stack are able to adequately represent the whole muscle volume.

A summation of CSA measurements from individual images is commonly used to estimate the whole muscle volume. It is generally assumed that individual CSA measurements are an appropriate surrogate for whole muscle volume measures. Some prior works have tested this assumption and have shown that individual CSA measurements are correlated with overall muscle volume in the quadriceps femoris (Morse *et al* 2007), vastii, adductors, hamstring (Mathur *et al* 2008), triceps surae (Albracht *et al* 2008) and anterior tibial (Esformes *et al* 2002, Lund *et al* 2002) muscle groups. Given this information, it would be useful to examine a wider range of lower-limb muscle groups to see whether it is possible to represent muscle volume by a smaller subset of CSA measurements and hence reduce operator time.

Typically, when prior studies have examined the correlation between the CSA and muscle volume measures, they are cited by others as validation that a particular measure can be used to assess the effect of an 'intervention', such as exercise, injury or disuse, on muscle (e.g. Esformes *et al* (2002) is cited by Reeves *et al* (2004) for this purpose). Whilst a significant correlation may be found between muscle volume and, for example, a CSA measure as part of a cross-sectional study, this does not imply that this same measure will be able to accurately

track changes in muscle volume as part of an interventional study. The goal of this study is to analyse different algorithms in their accuracy in predicting muscle volume changes due to an intervention based on a limited number of MRI images. One work (Tracy *et al* 2003) has considered this issue for the quadriceps muscle, but a similar analysis considering a number of muscles in the lower limb has not been conducted. In the case of this work, we draw on data considering changes in muscle volume during prolonged bed-rest with and without exercise in male individuals (Belavý *et al* 2009a, 2009b).

2. Materials and methods

2.1. Bed-rest study characteristics

Twenty male individuals underwent 56 days of strict bed-rest as part of the Berlin Bed-Rest Study. The study goal, methodology and subject inclusion and exclusion criteria have been described and discussed elsewhere (Rittweger *et al* 2006, 2010). Strict horizontal bed-rest was employed; subjects performed all hygiene in the supine position and were discouraged from moving excessively or unnecessarily. The institutional ethics committee approved this study and subjects gave their informed written consent. Subjects were randomized to either a countermeasure exercise group ($n = 10$) or an inactive control group ($n = 10$). The countermeasure exercise protocol, performed a total of 11 times per week, is discussed in detail elsewhere (Armbrecht *et al* 2010, Belavý *et al* 2009b). The effects of bed-rest (Belavý *et al* 2009a) and the exercise countermeasure (Belavý *et al* 2009b) on muscle volume changes during bed-rest have been published elsewhere.

2.2. Magnetic resonance imaging protocol, image analysis and data characteristics

Baseline MR scanning was conducted on the first day of bed-rest and then at 2 week intervals (day 14, day 28, day 42 and day 56) through to the end of the bed-rest period. Subjects were positioned on the scanning bed in supine with their knees and hips supported in slight flexion by a pillow under the knee. Transverse MR images were acquired using a 1.5 Tesla Magnetom Vision system (Siemens, Erlangen, Germany) from the lower limbs. Typically, 35 images of the thigh (from the superior aspect of the head of femur to the knee joint line; thickness = 10 mm; interslice distance = 5 mm, TR = 6000 ms, TE = 15 ms, FA = 180°, field of view: 480×480 mm interpolated to 512×512 pixels) and 30 images of the lower leg (knee joint line to the distal most portion of the lateral malleolus; thickness = 10 mm; interslice distance = 5 mm, TR = 4800 ms, TE = 15 ms, FA = 180°, field of view: 340×340 mm interpolated to 512×512 pixels) were acquired, though for taller subjects, additional images were added to ensure the region of interest was captured. Images were stored for offline analysis.

One operator (TM) performed all image measurements. Only the left leg was considered in analyses. To ensure operator blinding to study time-point and subject group, each data set was assigned a random number (www.random.org). ImageJ (Ver. 1.38x, http://rsb.info.nih.gov/ij/) was used for MR image analysis. The CSA of the following muscles in the thigh was measured in each image: rectus femoris, vastii, sartorius, gracilis, adductor magnus, adductor longus, biceps femoris long head, biceps femoris short head, semitendinosus and semimembranosus. Whilst adductor magnus and adductor longus could be readily differentiated from adductor brevis, this adductor muscle could not easily be differentiated from pectineus and hence was not measured. In the lower leg the following muscles were measured: gastrocnemius lateralis, gastrocnemius medialis, soleus with flexor hallucis longus (as too few anatomical landmarks (e.g. fascia) were available on MRI, soleus

was difficult to separate from flexor hallicus longus in a number of subjects), tibialis posterior, flexor digitorum longus, peroneal group (peroneus longus, brevis and tertius), anterior tibial muscles (tibialis anterior, extensor digitorum longus, extensor hallucis longus). More details regarding image measurement can be found in Belavý *et al* (2009a, 2009b). Muscle volume was calculated via linear interpolation (given a slice thickness of 10 mm and interslice distance of 5 mm).

A total of 90 lower leg and 91 thigh data sets were available for analysis. In some instances (such as due to scanner failure or movement artefacts; see Belavý *et al* (2009b)) data were not available from all subjects on all study dates.

2.3. Muscle volume prediction algorithms

As our interest was in change in muscle volume, we test prediction of changes in muscle volume due to bed-rest based on a limited number of MR images. Five different algorithms were tested in their accuracy to predict muscle volume based on a limited number of MRI slices. For each algorithm, muscle volume was derived from CSA measurements and compared with muscle volume derived from the entire MRI data set (figure 1).

- Algorithm 1: the largest single CSA measurement as well as the sum of the 3, 6, 9, 12, 15 and 18 largest CSA measurements. 7 sub-algorithms in total (1–18 CSA measurements in each muscle size estimate):

$$\hat{s} = \frac{\sum_{i=0}^{n} (A_{(P.i)})}{n}, \tag{1}$$

where $\hat{s}$ is the muscle size estimator (average CSA), A is the muscle CSA in a given image, $i = 0$ corresponds to the largest CSA measurement, and $i = 1$ corresponds to the image ($P.1$) with the second largest CSA measurement) and n is the total number of CSA measurements.

- Algorithm 2: the position of the largest CSA measurement (P.0) was located and taken as the first measurement of muscle size, then incorporating the next CSA immediately adjacent and distal to P.0 (2 slices). For the next sub-algorithm, the same two CSA measures were then also used with the third CSA measure immediately adjacent and proximal to P.0 (3 slices). This process of progressively adding one slice distally and then proximally was repeated until a total of 23 CSA measures were incorporated in the final sub-algorithm:

$$\hat{s} = \frac{A_{(P.0)} + \sum_{i=0}^{n_i} (A_{(P.0+i)}) + \sum_{k=0}^{n_k} (A_{(P.0-k)})}{1 + n_i + n_k} \tag{2}$$

where $\hat{s}$ is the muscle size estimator (average CSA), $A_{(P.0)}$ is the largest CSA measurement, i indicates images that are distal to the image with maximal muscle CSA ($P.0$), k indicates images that are proximal to the image with maximal muscle CSA ($P.0$), A is the CSA of the muscle in the image in question and n_i and n_k are the number images distal and proximal to image $P.0$. Here n_i and n_k increase asynchronously (first n_i by 1, then n_k by 1, etc). Where the proximal or distal end of the muscle was reached, further CSA measurements were included only from the end of the muscle that was still present.

- Algorithm 3: the same as algorithm 2 except that every second image was taken distally and proximally, thus covering a larger overall muscle volume with fewer individual CSA measurements (up to 19 slices).
 - 19 sub-algorithms in total (1–19 CSA measurements in each muscle size estimate):

$$\hat{s} = \frac{A_{(P.0)} + \sum_{i=0}^{n_i} (A_{(P.0+2i)}) + \sum_{k=0}^{n_k} (A_{(P.0-2k)})}{1 + n_i + n_k}. \tag{3}$$
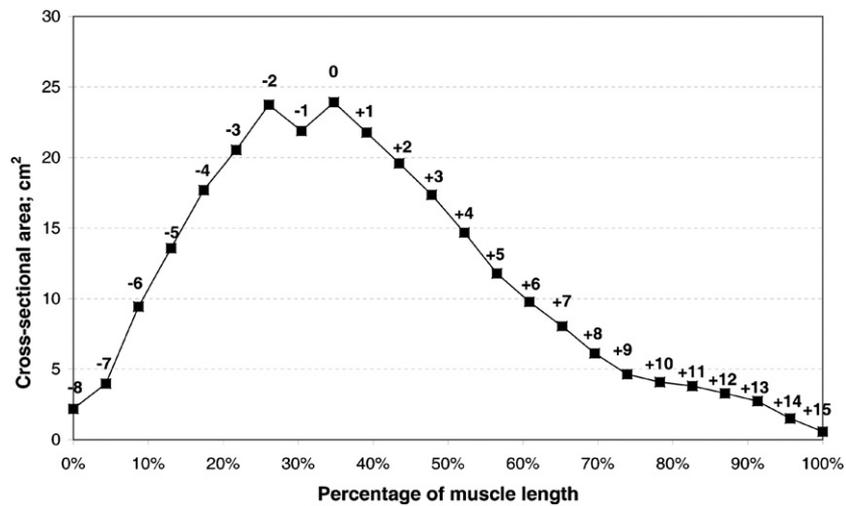
Figure 1. Description of muscle size measurement algorithms. Data presented are from the soleus muscle in one subject. '0' denotes the image with the largest CSA measure. Positions proximal and distal from image 0 are then given as negative and positive numbers, respectively. Algorithm 1 selects the largest CSA measurement (marked as image '0' in the plot) and then progressively adds more and more CSA measurements to the muscle size estimate (progressing to the smallest; i.e. in the above data set, in the order $0, -2, -1, +1, -3, +2, -4 \ldots$). Algorithm 2 selects the largest CSA measurement ('0' above), and then for the next sub-algorithm selects the next most distal image ('+1' above), then the next proximal image ('−1' above), and progressively including more and more images surrounding the peak CSA measurement to the measurement of muscle size (i.e. in the image order $0, +1, -1, +2, -2, +3 \ldots$). Algorithm 3 is the same as algorithm 2, except that every 2nd CSA measurement is skipped, so that the inter-image distance is 2 cm and a greater volume of muscle is covered (i.e. in the image order $0, +2, -2, +4, -4, +6, -6 \ldots$). Algorithm 4 selects images at 30% of muscle length (interpolated between images −2 and −1 above), 50% (interpolated between images +4 and +5 above), 80% (interpolated between images +10 and +11 above) as well as the sum of these three CSA measures. Algorithm 5: starting with the most proximal CSA measurement (−8 in the example above), every nth slice is measured (i.e. every 2nd: $-8, -6, -4, -2, \ldots, +12, +14$; every 3rd: $-8, -5, -2, +1 \ldots$) with the sub-algorithms evaluating up to every 16th slice.

- Algorithm 4: the CSA at 30%, 50% and 80% of muscle length was calculated. If these positions fell between two images, linear interpolation between the two adjacent images was used to calculate the CSA measure at each percentage of muscle length.

  - 4 sub-algorithms in total (30%, 50%, 80% and $\text{Sum}_{30,50,80\%}$).
  - Muscle length was defined by the number of images in which the muscle was present (n_{images}).
  - Where the 30%, 50%, or 80% lengths were not whole numbers, the CSA between the two adjacent images was linearly interpolated (e.g. when n_{images} was 14 with the 30% position corresponding to the '4.2th image', then $\text{CSA}_{30\%} = \text{CSA}_{\text{image4}} + (0.2 * (\text{CSA}_{\text{image5}} - \text{CSA}_{\text{image4}}))$.

- Algorithm 5: from the most proximal CSA measurement, sub-algorithms considered every 2nd, 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th, 11th, 12th, 13th, 14th, 15th or 16th CSA measurements in the calculation of average muscle CSA (i.e. a certain number of images were skipped between CSA measurements in the calculation of muscle size).
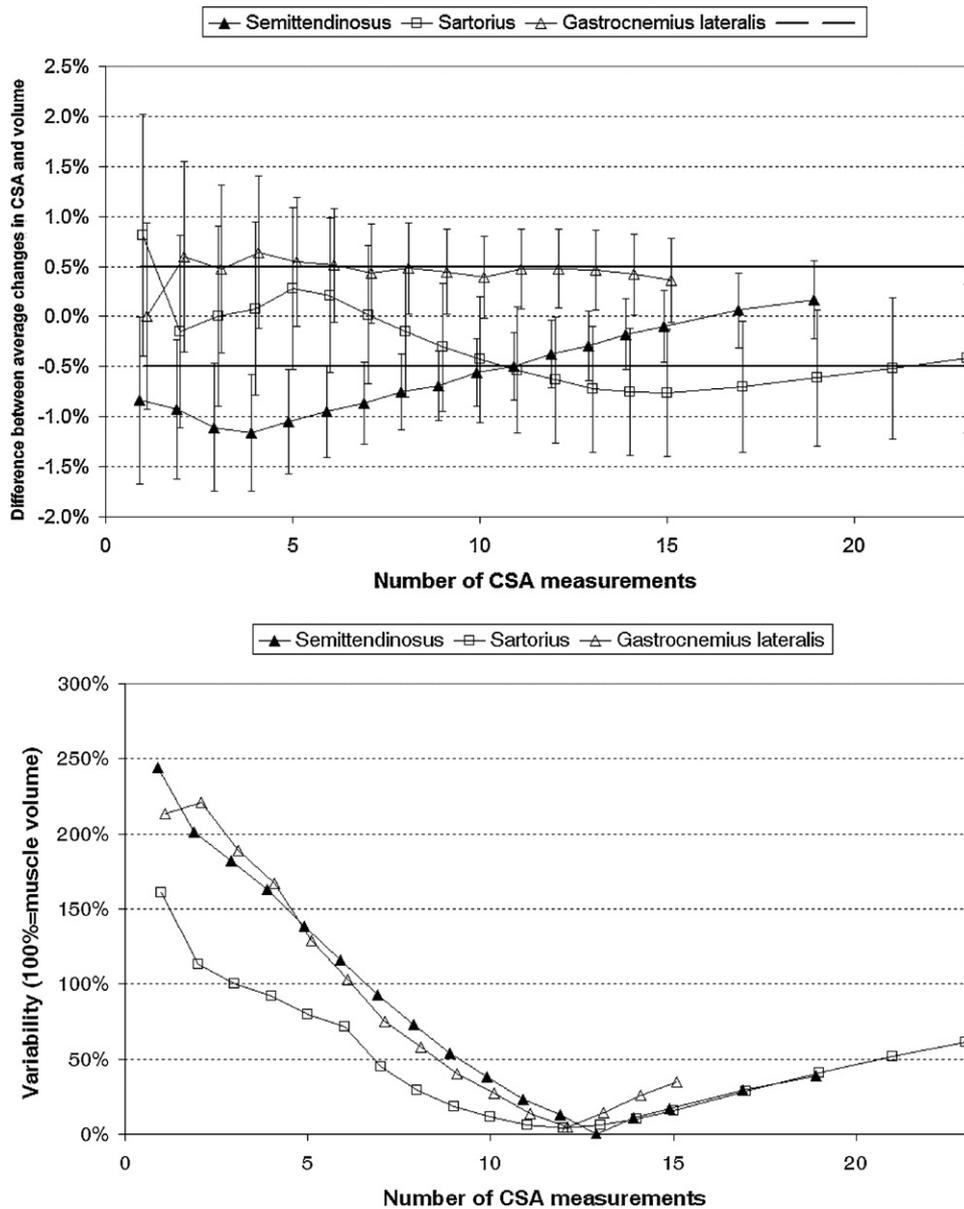
Figure 2. Example data from algorithm 2 testing for three muscles. Statistical modelling provided the difference in mean percentage change in muscle size during bed-rest for each sub-algorithm compared to muscle volume (top) as well as estimates of residual variance (bottom: these were then expressed as multiples of the variance when entire muscle volume was used (i.e. 100%)). Criteria 1 and 4 (top) were fulfilled with 11 CSA measurements in semitendinosus, 7 in gastrocnemius lateralis and 23 in sartorius. The variability of the percentage change in muscle size measures (criterion 2; bottom), obtained from statistical modelling, was higher than that of the percentage change in muscle volume with few CSA measures in the measure of muscle size, and this first reduced below that of muscle volume (100%) with four CSA measures in sartorius and seven in both gastrocnemius lateralis and semitendinosus. Thus, overall a minimum of 11 CSA measures are needed for semitendinosus, 23 for sartorius and 7 for gastrocnemius lateralis. The data presented in table 2 indicate the minimum number of CSA measurements to fulfil these criteria. See figure 3 and text for further details regarding variability.

For every measurement of muscle CSA calculated as part of each algorithm and sub-algorithm, the percentage change in relation to the baseline measurement (day 1 of bed-rest) for each muscle and subject was then calculated for each study date (days 14, 28, 42, 56) for use in further statistical analysis. In instances where baseline data were missing (see Belavý *et al* 2009b), the percentage change compared to day 14 was calculated on days 28, 42 and 56. Custom written software (in the Labview 6i environment, www.ni.com/labview) was used to perform this data processing.

2.4. Statistical analysis

The main goal of this paper is to examine the accuracy of a limited number of CSA measurements to provide an accurate estimation of changes in muscle volume. As correlation coefficients are typically insensitive to estimator bias, in order for an algorithm, or sub-algorithm, to be considered accurate, the following criteria were used.

- Criterion 1: the mean percentage change in the muscle CSA needed to be within 0.5% of the mean percentage change in muscle volume. See also figure 2.
- Criterion 2: the variability of the percentage change in muscle size was required to be the same or less than that of the variability of the percentage change in muscle volume. As part of statistical modelling using linear mixed-effects models, allowances for heterogeneity of variance (in percentage change in muscle volume) according to 'sub-algorithm' were made and the variance estimates were obtained from statistical modelling. The rationale for this criterion was that if the variability of the percentage change in muscle size were higher than that of muscle volume then this would indicate a greater degree of measurement error and hence, compared to muscle volume, restricted ability to detect the effects of an intervention. See also figure 2 (bottom panel) and figure 3 for further details.
- Criterion 3 (for algorithms 1, 2 and 3): a sub-algorithm was also rejected if the muscle length encompassed by the number of images was greater than the median muscle length. This median muscle length was derived from the available dataset of 20 subjects. This would otherwise imply that the amount of muscle volume required in order for the algorithm to succeed would be greater than that present.
- Criterion 4 (for algorithms 1, 2 and 3): sub-algorithm was also rejected if increasing the number of CSA measurements further resulted in criteria 1 and 2 being rejected (e.g. if criteria 1 and 2 were met with one CSA measure only, but not with two to seven CSA measures, then sub-algorithm with one CSA would be rejected).

These criteria were chosen as they are independent of the actual extent of change in muscle volume during bed-rest with or without exercise, and they focus on the accuracy of predicting the muscle volume change, whatever this may be. Linear mixed-effects models (Pinheiro and Bates, 2000) in the 'R' statistical environment (version 2.4.1, www.r-project.org) were used to generate the data to assess criteria 1 and 2. Fixed effects of 'day of bed-rest' (day 14, day 28, day 42, day 56) and 'sub-algorithm' (e.g. 30%, 50%, 80%, $Sum_{30,50,80\%}$ and muscle volume for testing algorithm 4) as well as their two-way interaction were used. Allowances for heterogeneity of variance for 'day of bed-rest' were made. To test whether the variability of the percentage change in muscle size (criterion 2) differed across sub-algorithms and muscle volume measures, a second model was fitted with allowances for heterogeneity of variance for 'sub-algorithm' as well. The variability estimates were then obtained from the model to assess whether the variability of the percentage change in muscle volume of a given sub-algorithm was higher or lower than that of muscle volume. The difference in mean percentage
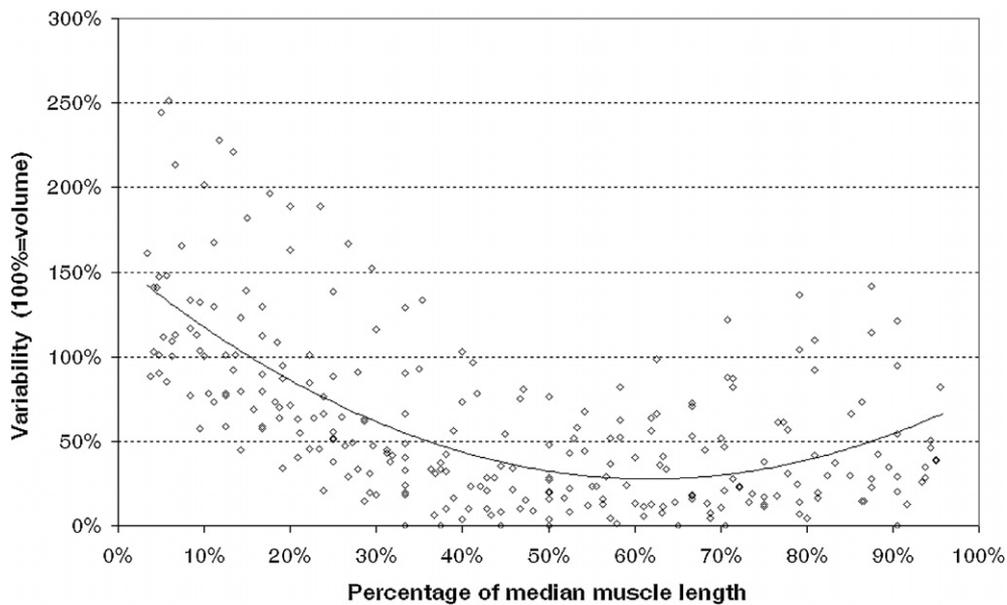
Figure 3. Data here from algorithm 2 show that variability of the percentage change in muscle size is less when only part of the muscle is measured. Estimates of residual variance were obtained from statistical modelling and these were then expressed as multiples of the variance when entire muscle volume was used (i.e. 100%). The number of CSA measures has been expressed as a percentage of the median number of CSA measures per muscle (i.e. as a percentage of muscle length). The variability of the muscle size measurement, obtained from statistical modelling, was high when few CSA measures were included. As more CSA measures were included, variability decreased and reached lower levels (∼30%). However, once more than 60–65% of the CSA measures were included, variability again increased, such that measuring entire muscle volume produced greater variability compared to measuring ∼60–65% of the muscle volume. This may seem counter-intuitive, but was probably due to the fact that as the last few CSA measurements were included (i.e. using algorithm 2, the anatomical origin and insertion of the muscle were reached), CSA is typically lower. Typically, operator measurement error will not decrease to the same extent, leading to a relative increase in variability. These data suggest that including 60–65% of a muscle's length in a measurement result in lowest outcome measure variability.

change in muscle size compared to muscle volume was then calculated for each sub-algorithm (criterion 1). Pearson's correlation coefficient was also calculated between the muscle volume and the raw measures of muscle size from each sub-algorithm.

3. Results

For brevity, details of algorithm testing are not presented here but are available in the online supplementary material from stacks.iop.org/PM/32/35/mmedia (tables 4–7). For descriptive purposes, the results of algorithm 2 for three muscles are presented in figure 2. A summary of the results of algorithm testing is presented in table 1 along with data on muscle dimensions (volume, length). The table provides information on the minimum number of CSA measurements required for an acceptable estimate of muscle volume (for algorithms 1, 2 and 3) and whether measures taken at 30%, 50% or 80% of muscle length (algorithm 4) or by

skipping a pre-defined number of slices (irrespective of the location of the peak CSA; algorithm 5) are appropriate for muscle volume estimates. For a typical data set (measurements from all of the muscles on one lower limb) there were 366 individual CSA measurements (representing the sum of the median image numbers in table 1).

Using either algorithms 1, 2 or 3, typically required 193, 191 or 232 individual CSA measurements, respectively, to accurately estimate the volumes of all lower-limb muscles considered. This summary result indicated that algorithms 1 and 2 were generally better than algorithm 3. A comparison of algorithms 2 and 3 revealed that algorithm 2 resulted in the lowest numbers of required CSA measurements required to accurately estimate muscle volume for all muscles except the vastii, peroneals, sartorius and anterior tibial muscles. For these muscles as indicated by algorithm 3, selecting images spaced every 2 cm (thus covering a larger muscle volume) further reduced the number of images required for analysis. For both heads of biceps femoris in addition to the semimembranosus and gracilis, typically the whole muscle volume needed to be measured.

For algorithm 4, in no case did the approach of summating the CSA at 30%, 50% and 80% of muscle length (Sum$_{30,50,80\%}$) fits the criteria for acceptance. In some muscles (anterior tibial, gastrocnemius medialis, peroneal, tibialis posterior, thigh muscles, adductor longus, adductor magnus, rectus femoris, sartorius) selecting single CSA measures at 30%, 50% or 80% of muscle length could represent appropriate estimates of muscle volume.

Only for the adductor longus muscle did algorithm 5 satisfy the acceptance criteria: the sub-algorithm of choosing every second slice (starting with the most proximal CSA measurement) could be used as a surrogate measure, which was comparable (in terms of the number of CSA measurements needed) to the results of algorithm 2 for this muscle. For all other muscles, none of the sub-algorithms for algorithm 5 satisfied acceptance criteria.

Correlations between the raw measures of muscle size and the gold standard of muscle volume are presented for select algorithms in table 2. (Data on correlations for remaining muscle size measurement approaches are available from authors upon request.) As could be expected, as the number of slices included in the measure of muscle size increased (algorithm 2), the correlation with muscle volume tended to approach '1' (which would indicate a 'perfect' correlation). All the correlations were statistically significant ($p < 0.01$). Data on muscle volume prediction are contained in table 3.

Variability of the percentage change in muscle size measures (used in assessing criterion 2) decreased as more CSA measures were included and also decreased below that of the muscle volume measure (figure 3). Fitting a quadratic model to these data ($R^2 = 0.58$) indicated that minimum variability was achieved when approximately 60–65% of total muscle length was captured around the image with the peak CSA. At this point, estimates from quadratic fit indicated that variability would be approximately 30% of that when the entire muscle volume is measured, though this varied from muscle to muscle.

4. Discussion

As a large amount of operator time is required to manually measure muscle volume from MR images, in the current work we aimed to evaluate the ability of certain algorithms to provide appropriate surrogate measures of changes in muscle volume, based upon a limited number of MR images, associated with an 'intervention': in this case bed-rest with and without exercise. Whilst one work has examined this issue for the quadriceps muscle (Tracy *et al* 2003), we wished to provide useful research tools for assessing changes in muscle size in a wide variety

Table 1. Muscle dimensions and results of algorithm testing.

Muscle	Muscle characteristics				Results of algorithm testing				
	Volume mean (SD) cm^3	Mean(SD) change in volume after 56 days bed-rest	Median length (min−max) number of images	Mean length (in cm)	1	2	3	4	5
Calf muscles									
Anterior tibial muscles	253.6(20.9)	−4.9(4.5)%	24(21–27)	36.0	6	14	10	30%, 50%	–
Flexor digitorum longus	32.7(7.3)	−3.3(10.3)%	17(15–20)	25.8	12	7	17*	–	–
Gastrocnemius lateralis	157.0(34.7)	−6.1(10.9)%	15(11–18)	22.1	15*	7	7	–	–
Gastrocnemius medialis	231.4(57.0)	−14.7(8.5)%	16(13–20)	25.0	3	2	5	80%	–
Peroneal	142.9(25.0)	−5.9(8.4)%	24(20–27)	36.2	9	13	7	80%	–
Soleus	490.3(77.7)	−12.1(6.8)%	22(17–26)	33.4	18	4	10	–	–
Tibialis posterior	110.4(22.5)	−6.4(7.9)%	21(19–25)	32.0	3	12	21*	80%	–
Thigh muscles									
Adductor longus	183.6(35.2)	−2.2(7.4)%	12(10–15)	18.8	12*	3	12*	50%	2nd
Adductor magnus	560.2(78.6)	−6.0(5.4)%	18(14–25)	26.2	9	1	18*	50%	–
Biceps femoris long head	212.2(34.6)	−8.8(8.6)%	19(13–24)	28.1	15	19*	19*	–	–
Biceps femoris short head	115.5(25.9)	−2.6(9.1)%	16(13–21)	24.4	9	16*	16*	–	–
Gracilis	112.0(19.8)	−1.4(6.7)%	21(17–26)	31.5	21*	21*	21*	–	–
Rectus femoris	312.5(51.2)	−0.6(7.5)%	21(17–25)	31.9	3	9	21*	50%, 80%	–
Sartorius	164.5(30.6)	−2.4(5.7)%	30(27–33)	45.5	30*	23	7	50%, 80%	–
Semimembranosus	246.4(41.5)	−8.4(9.3)%	18(12–21)	25.5	15	18*	18*	–	–
Semitendinosus	227.4(35.4)	−5.9(8.1)%	20(17–27)	30.9	12	11	20*	–	–
Vastii	1747.6(188.6)	−8.6(6.8)%	27(24–29)	39.6	1	11	3	–	–

Muscle dimensions were averaged across all scanning dates for each subject and then across all subjects. The second column from left indicates the number of images in which the muscle was present. The mean muscle length was calculated given a slice thickness of 1 cm and interslice distance of 0.5 cm (i.e. mean number of images * 1.5 cm). In results of modelling, '*' and '–' indicate that no sub-algorithm fits the acceptance criteria (see section 2.4) and that either a different algorithm should be chosen or the entire muscle volume measured. For '*', the number given represents the median number of CSA measures for this muscle. Data on percentage change in muscle volume are pooled from all bed-rest subjects, and detailed effects of bed-rest with and without exercise are reported and discussed elsewhere (Belavý et al 2009a, 2009b).

Table 2. Correlations between muscle volume and muscle size for algorithms 2 and 4.

Muscle	Algorithm 2: number of contiguous CSA measurements including largest CSA								Algorithm 4: CSA at percentage of muscle length			
	1	2	3	5	7	9	12	14	30%	50%	80%	Sum$_{30, 50, 80\%}$
Calf muscles												
Anterior tibial muscles	0.52	**0.53**	**0.52**	**0.57**	**0.64**	0.71	0.81	**0.86**	**0.45**	**0.67**	0.64	0.73
Flexor digitorum longus	0.91	0.91	0.92	0.93	**0.95**	**0.97**	**0.98**	**0.98**	0.80	0.85	0.60	0.93
Gastrocnemius lateralis	0.84	0.86	0.86	0.91	**0.93**	**0.95**	**0.98**	**0.99**	0.77	0.83	0.65	0.91
Gastrocnemius medialis	0.91	**0.91**	**0.91**	**0.93**	**0.94**	**0.96**	**0.98**	**0.99**	0.91	0.89	**0.53**	0.89
Peroneal	0.82	0.81	0.81	0.84	0.87	0.89	0.90	**0.91**	0.78	0.91	**0.66**	0.91
Soleus	0.88	0.89	0.89	**0.90**	**0.91**	**0.93**	**0.96**	**0.97**	0.78	0.67	0.56	0.85
Tibialis posterior	0.88	0.90	0.91	0.91	0.91	0.90	**0.90**	**0.88**	0.89	0.79	**0.65**	0.91
Thigh muscles												
Adductor longus	0.91	0.91	**0.92**	**0.94**	**0.95**	**0.97**	**1.00**	**1.00**	0.86	**0.88**	0.49	0.89
Adductor magnus	**0.69**	**0.71**	**0.71**	**0.75**	**0.80**	**0.84**	**0.90**	**0.96**	0.68	**0.58**	0.21	0.63
Biceps femoris long head	0.68	0.69	0.71	0.76	0.81	0.87	0.91	0.95	0.48	0.69	0.43	0.73
Biceps femoris short head	0.92	0.91	0.92	0.93	0.94	0.96	0.97	0.98	0.77	0.88	0.77	0.90
Gracilis	0.88	0.90	0.91	**0.91**	0.90	0.91	0.91	0.92	0.88	0.90	0.55	0.91
Rectus femoris	0.79	0.80	0.80	0.82	0.83	**0.85**	**0.88**	**0.91**	0.79	**0.77**	**0.75**	0.83
Sartorius	0.86	0.91	0.91	0.92	0.90	0.91	0.93	0.94	0.86	**0.86**	**0.91**	0.94
Semimembranosus	0.81	0.79	0.82	0.82	0.81	0.81	0.82	0.86	0.64	0.76	0.80	0.84
Semitendinosus	0.84	0.84	0.84	0.84	0.84	0.85	**0.87**	**0.89**	0.83	0.74	0.46	0.83
Vastii	0.80	0.63	0.71	0.76	0.79	0.82	**0.85**	**0.88**	0.79	0.79	0.82	0.86

Values are Pearson's correlation coefficient; p all <0.01. Boldface numbers indicate that the sub-algorithm fit the acceptance criteria 1 and 2 (see section 4.2 and table 1). Note that although certain measures of muscle size may be highly, and significant statistically, correlated with the gold standard of muscle volume (such as biceps femoris short head, algorithm 2, 12 CSA measures), there is no guarantee that such measures are appropriate estimates of muscle volume.

Table 3. Prediction of muscle volume from average CSA measures.

Muscle	Algorithm	Sub-algorithm (number of CSA measures)	Average CSA (mm^2)			Parameters of fit		
			Minimum	Mean	Maximum	a	b	R^2
Calf muscles								
Anterior tibial muscles	3	10	604.0	783.4	944.4	0.269(0.017)‡	42.2(13.7)†	0.94
Flexor digitorum longus	2	7	113.0	193.0	290.1	0.151(0.007)‡	3.4(1.4)*	0.99
Gastrocnemius lateralis	2	7	656.9	975.1	1408.2	0.124(0.008)‡	35.1(8.4)‡	0.97
Gastrocnemius medialis	2	2	774.8	1463.0	2299.0	0.157(0.007)‡	1.1(11.4)	0.99
Peroneal	3	7	290.5	449.4	755.8	0.137(0.015)‡	81.3(7.3)‡	0.96
Soleus	2	4	1767.2	2632.5	3391.7	0.164(0.007)‡	55.4(21.1)*	0.98
Tibialis posterior	2	12	243.6	411.1	647.9	0.141(0.014)‡	52.7(6.6)‡	0.98
Thigh muscles								
Adductor longus	2	3	1112.1	1567.1	2165.9	0.101(0.009)‡	24.4(14.5)	0.96
Adductor magnus	2	1	2944.3	3669.9	5171.5	0.109(0.013)‡	160.3(50.7)†	0.94
Biceps femoris long head	Volume	All						
Biceps femoris short head	Volume	All						
Gracilis	Volume	All						
Rectus femoris	2	9	939.7	1358.8	1871.0	0.122(0.016)‡	145.9(22.8)‡	0.96
Sartorius	3	7	300.3	400.9	616.7	0.305(0.027)‡	42.2(11.2)‡	0.96
Semimembranosus	Volume	All						
Semitendinosus	2	11	668.5	992.2	1434.1	0.168(0.014)‡	60.7(14.6)‡	0.96
Vastii	3	3	5182.6	6371.7	8001.3	0.265(0.014)‡	55.8(94.6)	0.97

Volume in cm^3. CSA in mm^2: total area in all images measured divided by number of images measured. Prediction assumes the form: volume (cm^3) = (**a***Average CSA (mm^2)) + **b**. * $p < 0.05$; $^†p < 0.01$; $^‡ p < 0.001$. The coefficient of determination (R^2) is given for each prediction equation. Assuming $b = 0$ led to similar coefficients of determination for the fitted models. If these equations are used to predict muscle volume, results may be less accurate if average CSA falls outside the minimum/maximum range given. Where muscle volume could not be appropriately estimated by an algorithm, no data are given. The algorithms presented represent the 'best' (met criteria + least number of CSA measures) algorithms for a given muscle.

of lower-limb muscles. The main findings of the study were algorithms for predicting muscle volume and muscle volume change based upon a limited number of MR images. Overall, however, the results of algorithm testing in the current work suggest that using an anatomy-based approach (such as by orienting to areas of the maximal CSA) would generally be more successful than arbitrarily choosing select sections of a muscle. Nonetheless, the findings of the current study also suggest that there is no 'one size fits all' algorithm that is best for estimating muscle volume in all the musculature of the lower limb: the optimal algorithm will vary depending upon the anatomy of a muscle.

In considering the individual algorithms, the first algorithm we trialled (algorithm 1) took the largest CSA measure for a particular muscle and successively added the largest CSA measures in descending order, down to the smallest value. Whilst the results suggest that for some muscles, algorithm 1 requires fewer CSA measurements than other approaches; we would not recommend this approach as there is no guarantee that the largest CSA measures are necessarily located in adjacent images (which is particularly the case for long strap-like muscles such as sartorius). Hence, the operator would have the additional difficulty of locating the images with the largest CSAs, necessitating more measurements than indicated by algorithm testing. We would advise against its use and concentrate therefore on the remaining algorithms.

The results of overall algorithm testing showed that algorithm 2, measuring the image with the greatest muscle CSA and then progressively adding measurements immediately distal and proximal to this, provided the most efficient and efficacious reduction in measurement time. The results showed that $\sim$370 individual CSA measures are required to measure muscle volume of one lower limb in a typical male subject (assuming a MR-slice thickness of 1 cm and interslice distance of 0.5 cm). These results suggest that using algorithm 2 would almost halve the required number of measurements to accurately estimate muscle volume. Using this technique, however, the total number of CSA measurements required varied from one muscle to the next.

Whilst algorithm 2 was the optimal surrogate measure for muscle volume for the majority of lower-limb muscles, there were a few exceptions. For the vastii, peroneal, sartorius and anterior tibial muscle groups, analysis showed that it would be more time efficient, if the operator were to select images spaced wider apart (2 cm; algorithm 3), thus assessing a greater muscle volume with fewer measurements. For both heads of biceps femoris, semimembranosus and gracilis, analysis showed that it is not possible to select a smaller number of CSA measures and the entire muscle volume must be measured. Some caution should be applied for adductor magnus, as it is likely that a proximal part of the muscle was not captured in the imaging process (see supplementary data available at stacks.iop.org/PM/32/35/mmedia, figure 5). For the remaining eight muscles groups, algorithm 2 remained the most efficient approach, with the necessary number of measurements listed in table 1.

Algorithm 4 considered whether CSA measures taken along a muscle (30%, 50% and 80% of muscle length as well as the sum of these three measures) could provide appropriate estimates of muscle volume. In no case did averaging these three CSA measures prove accurate in tracking changes of muscle volume due to the intervention of bed-rest. In only a limited number of muscles did testing suggest that choosing a CSA measurement at a particular length of the muscle could provide an appropriate surrogate measure for muscle volume. In a number of these positive cases, the correlation between raw muscle volume and the CSA at the particular muscle length was not very high, typically being in the range of 0.45–0.77. As the positions along the length of the muscles were chosen arbitrarily, and not on anatomical considerations (such as in relation to the region of peak CSA), we would

advise further validation of the successful measures from this algorithm with another unrelated dataset before these approaches are implemented in future work. Such further testing would ensure that these results could be generalized to other datasets and were not just peculiar to the dataset evaluated in the current work.

Algorithm 5 considered whether starting at the most proximal point of a muscle and then skipping a fixed number of images between CSA measures (e.g. to choose every second or fourth image as in Tracy *et al* (2003)) could be used instead of having to measure every image. Based on the criteria we used for accepting algorithms, only in one case (adductor longus, every 2nd CSA measure) was this approach successful. This approach failed mainly because the sub-algorithms variability was almost always higher than the muscle volume measure and researchers would need to be satisfied with greater outcome measure variability in order to be able to implement this approach (see also supplementary table 7 available at stacks.iop.org/PM/32/35/mmedia). It may at first seem surprising that algorithm 2 requires only one CSA measurement for adductor magnus but that algorithm 5 requires many more. It should be remembered that the two algorithms have different anatomical starting points and different progressions (in terms of where additional slices come from). For these reasons it is understandable why they yield different results.

An important secondary finding of the current study was that those algorithms found to be accurate in tracking changes in muscle volume were also very good predictors of raw muscle volume, when considering the correlations between muscle volume and muscle size (table 2). The converse of this finding was that although a correlation between muscle volume and muscle size measure from one of the algorithms may have been 'high', this did not imply that the same measure could accurately track changes in muscle volume during the intervention of bed-rest. Based upon data presented in other works (Albracht *et al* 2008, Esformes *et al* 2002, Lund *et al* 2002, Mathur *et al* 2008, Morse *et al* 2007) one may be temped to argue that a 'significant correlation' between raw muscle volume and another measure of muscle size (e.g. one CSA measurement) indicates that the surrogate muscle CSA measure can be used instead of muscle volume in an interventional study. The results of the current study clearly show that this assumption cannot be made. It is important to remember that a correlation coefficient (r) of 0.5 means, in the current context of muscle size quantification, that only 25% of the variance in the muscle volume can be explained by the surrogate muscle size measure (see Moore and McCabe (2006), pp 141–2). Similarly an r of 0.9 still leaves 19% of the variability of the surrogate measure unrelated to muscle volume changes, which could be associated with greater measurement error and/or differences in muscle morphology. This is certainly the underlying reason for our findings that high and statistically significant correlations between muscle size measures and the whole muscle volume may be observed, but that the particular muscle size measure is inadequate for following the changes in muscle volume with an intervention.

Will the approaches suggested in the current study actually save operator time? If an operator, using MR images of 1 cm thickness at 0.5 cm separation, wished to measure all the lower-limb muscles evaluated in the current study, approximately 370 manual CSA measurements would be required to measure all muscle volumes. If the operator were to measure, as suggested by the results of the current study, the vastii, peroneal, sartorius and anterior tibial muscles using algorithm 3, the entire volume of the two heads of biceps femoris, semimembranosus and gracilis and the remaining nine muscles as indicated by algorithm 2, then ∼157 individual manual CSA measures would be required. This represents a reduction of ∼60% in the total number of image measurements. As an operator can visually assess the region of muscle with the greatest CSA and hence begin measuring here, the overall operator time saved will be considerable.

It is also important to consider the limitations of the findings from the current study. One issue that should be considered first is that we chose specific criteria for accepting algorithms. On the one hand, expecting that variability of the measure (criterion 2) to be at least the same as or less than the 'gold standard' change in muscle volume measure is a reasonable expectation. Choosing a threshold of 0.5% for the mean bias (criterion 1) could however be considered arbitrary. Other researchers may wish higher (or be satisfied with lower) accuracy. To this end, the data presented in the supplementary tables (available from stacks.iop.org/PM/32/35/mmedia) will enable other researchers to define algorithms based upon their preferences for 'accuracy'.

It should also be noted that the most successful algorithms (2 and 3) depend upon the detection, by the operator, of the MR image(s) with the maximal muscle CSA. Whilst in our experience this is not difficult for the skilled operator, it is imperative that those implementing such measurement approaches have a good knowledge of functional anatomy. Such tasks should not be left to the novice operator.

Furthermore, in our MR data we did not have contiguous slices, but rather had a 0.5 cm interslice gap with the 1 cm thick slices. Whilst this is unlikely to greatly influence the muscle volume estimates in the current work, it is likely that the variability of muscle volume measures would be even lower with contiguous slices (cf variability results from testing of algorithm 5). Related to this, we used linear interpolation for calculating muscle volume. It may be possible that more accurate interpolation methods could be used for calculating muscle volume (e.g. as presented in Nordez *et al* (2009)), though such analyses are beyond the scope of the current work.

As not all future studies will use the same MR-imaging approach (1 cm slice thickness, 0.5 cm interslice distance), for those muscle volumes that were suitably represented using algorithm 2, we would advise other researchers to measure an equivalent amount of muscle around the region of greatest CSA: for example, for soleus, four images, or 6 cm of muscle length using the current protocol would be required. For another imaging protocol, such as with 0.5 cm thick slices and 0.5 cm interslice distance, an equivalent amount of muscle would be captured by six images around the image with greatest CSA. The same methods can be applied for volume estimations of the muscle groups suitably represented by algorithm 3, by simply doubling the muscle length incorporated.

5. Conclusions

We examined the ability of different muscle size measurement approaches to accurately track changes in muscle volume as a result of bed-rest. We found that an approach selecting the MR image with the highest muscle CSA and then, depending on muscle, a number of CSA measures immediately distal and proximal, provided an acceptable estimate of muscle volume. In the vastii, peroneal, sartorius and anterior tibial muscle groups, results can be attained more time-efficiently by increasing the spacing between CSA measures, thus incorporating a larger muscle volume. In the two heads of biceps femoris, semimembranosus and gracilis, it is not possible to reduce the number of CSA measures and entire muscle volume must be evaluated. These findings help in choosing optimal algorithms for measuring muscle volume changes based on MRI.

Acknowledgments

References

Albracht K, Arampatzis A and Baltzopoulos V 2008 Assessment of muscle volume and physiological cross-sectional area of the human triceps surae muscle *in vivo J. Biomech.* **41** 2211–8

Armbrecht G, Belavý D L, Gast G, Bongrazio M, Touby F, Beller G, Roth H J, Perschel F H, Rittweger J and Felsenberg D 2010 Resistive vibration exercise attenuates bone and muscle atrophy in 56 days of bed rest: biochemical markers of bone metabolism *Osteoporos. Int.* **21** 597–607

Belavý D L, Miokovic T, Armbrecht G, Richardson C A, Rittweger J and Felsenberg D 2009a Differential atrophy of the lower-limb musculature during prolonged bed-rest *Eur. J. Appl. Physiol.* **107** 489–99

Belavý D L, Miokovic T, Armbrecht G, Rittweger J and Felsenberg D 2009b Resistive vibration exercise reduces lower limb muscle atrophy during 56-day bed-rest *J. Musculoskelet. Neuronal. Interact.* **9** 225–35

Blazevich A J, Coleman D R, Horne S and Cannavan D 2009 Anatomical predictors of maximum isometric and concentric knee extensor moment *Eur. J. Appl. Physiol.* **105** 869–78

Brem M H *et al* 2009 Magnetic resonance image segmentation using semi-automated software for quantification of knee articular cartilage—initial evaluation of a technique for paired scans *Skeletal Radiol.* **38** 505–11

Esformes J I, Narici M V and Maganaris C N 2002 Measurement of human muscle volume using ultrasonography *Eur. J. Appl. Physiol.* **87** 90–2

Fukunaga T, Miyatani M, Tachi M, Kouzaki M, Kawakami Y and Kanehisa H 2001 Muscle volume is a major determinant of joint torque in humans *Acta Physiol. Scand.* **172** 249–55

Lund H, Christensen L, Savnik A, Boesen J, Danneskiold-Samsoe B and Bliddal H 2002 Volume estimation of extensor muscles of the lower leg based on MR imaging *Eur. Radiol.* **12** 2982–7

Mathur S, Takai K P, Macintyre D L and Reid D 2008 Estimation of thigh muscle mass with magnetic resonance imaging in older adults and people with chronic obstructive pulmonary disease *Phys. Ther.* **88** 219–30

Moore D S and McCabe G P 2006 *Introduction to The Practice of Statistics* (New York: Freeman)

Morse C I, Degens H and Jones D A 2007 The validity of estimating quadriceps volume from single MRI cross-sections in young men *Eur. J. Appl. Physiol.* **100** 267–74

Narici M V, Roi G S, Landoni L, Minetti A E and Cerretelli P 1989 Changes in force, cross-sectional area and neural activation during strength training and detraining of the human quadriceps *Eur. J. Appl. Physiol. Occup. Physiol.* **59** 310–9

Nordez A, Jolivet E, Südhoff I, Bonneau D, de Guise J A and Skalli W 2009 Comparison of methods to assess quadriceps muscle volume using magnetic resonance imaging *J. Magn. Reson. Imaging* **30** 1116–23

Pinheiro J C and Bates D M 2000 *Mixed-Effects Models in S and S-PLUS* (Berlin: Springer)

Reeves N D, Narici M V and Maganaris C N 2004 Effect of resistance training on skeletal muscle-specific force in elderly humans *J. Appl. Physiol.* **96** 885–92

Rittweger J *et al* 2006 Highly demanding resistive vibration exercise program is tolerated during 56 days of strict bed-rest *Int. J. Sport. Med.* **27** 553–9

Rittweger J *et al* 2010 Prevention of bone loss during 56 days of strict bed rest by side-alternating resistive vibration exercise *Bone* **46** 137–47

Tracy B L, Ivey F M, Jeffrey Metter E, Fleg J L, Siegel E L and Hurley B F 2003 A more efficient magnetic resonance imaging-based strategy for measuring quadriceps muscle volume *Med. Sci. Sports Exerc.* **35** 425–33

Trappe S W, Trappe T A, Lee G A and Costill D L 2001 Calf muscle strength in humans *Int. J. Sports Med.* **22** 186–91

Supplementary Table 4: Results of algorithm 1 and algorithm 4 testing

Algorithm	Sub-algorithm	Muscle																
		V	SEMI_T	SEMI_M	SART	RF	GRAC	BFL	BFS	AM	AL	FDL	GLAT	GMED	PER	SOL	ANT	TIBP
		For criterion 1: Difference between percentage change in muscle size calculated by sub-algorithm versus percentage change in muscle volume (%)																
1	1	-0.119	-0.511	2.156	0.620	-0.175	0.742	-2.098	0.930	0.149	-0.573	-0.421	0.223	0.557	-2.247	-0.345	-0.679	-0.686
	3	-0.075	-0.725	2.329	-0.200	-0.065	0.872	-1.763	1.143	-0.105	-0.411	0.272	0.702	0.064	-1.019	0.170	-0.550	0.153
	6	-0.129	-0.640	2.246	-0.702	0.003	0.900	-0.718	1.540	-0.787	-0.603	0.197	0.400	-0.097	-0.540	0.447	-0.247	0.284
	9	-0.120	-0.672	1.998	-0.751	-0.236	1.026	-0.473	1.348	-0.198	-0.736	-0.052	0.095	0.262	-0.373	0.547	0.045	0.458
	12	-0.186	-0.381	1.356	-0.809	-0.254	1.082	-0.174	0.814	0.302	-0.025	0.121	-0.495	0.005	-0.270	0.597	0.015	0.310
	15	-0.166	-0.043	0.379	-0.937	-0.144	1.056	0.082	0.377	0.365	0.000	0.322	-0.173	-0.177	-0.257	0.504	0.000	-0.013
	18	-0.169	-0.007	-0.076	-1.043	0.137	0.675	0.077	0.133	0.110		-0.052		-0.035	-0.154	0.300	0.062	-0.013
4	30%	-2.268	-2.477	-1.562	0.690	0.520	1.025	1.167	1.879	-1.410	1.997	-2.539	-3.237	-1.248	-2.028	1.771	-0.254	0.838
	50%	-1.088	-0.828	0.322	-0.030	0.087	1.589	5.841	2.248	-0.278	0.077	-1.320	-0.451	-0.609	-0.714	1.208	0.495	1.248
	80%	-0.520	0.702	1.104	-0.046	0.458	1.009	0.950	1.638	0.351	-0.741	-0.216	-1.326	0.171	-0.409	1.322	0.756	0.094
	Sum$_{30,50,80\%}$	-0.503	0.652	0.921	-0.175	0.420	0.951	0.854	1.521	0.546	-0.981	-0.250	-1.280	0.155	-0.409	1.354	0.745	0.033
		For criterion 2: Variability (expressed as a fraction of the variability of the percentage change in muscle volume [=1])																
1	1	0.525	2.164	0.668	1.208	2.157	1.883	1.558	1.537	1.433	0.968	4.864	18.072	1.053	2.186	4.325	1.104	1.135
	3	0.377	1.622	0.496	0.819	1.160	1.222	1.209	1.114	1.116	0.748	3.895	15.110	0.845	1.385	3.762	0.819	0.714
	6	0.271	1.034	0.309	0.557	0.795	0.478	0.712	0.878	0.978	0.603	2.386	9.656	0.599	0.967	3.214	0.593	0.560
	9	0.192	0.515	0.137	0.368	0.420	0.198	0.529	0.514	0.764	0.425	1.128	6.312	0.414	0.421	2.318	0.273	0.338
	12	0.006	0.137	0.544	0.137	0.212	0.001	0.249	0.004	0.146	0.871	0.603	3.018	0.455	0.203	1.230	0.162	0.152
	15	0.160	0.335	0.847	0.189	0.324	0.172	0.617	0.518	0.649	1.000	0.541	0.001	0.851	0.284	0.474	0.509	0.366
	18	0.264	0.589	0.984	0.314	0.434	0.317	0.959	0.811	0.910		0.869		0.966	0.491	0.635	0.659	0.762
4	30%	1.868	1.504	3.602	1.640	1.241	1.363	1.978	2.180	2.478	2.190	3.231	4.116	3.147	1.324	1.787	0.871	1.007
	50%	1.443	1.592	0.833	1.770	802.990	1.643	0.288	0.763	1.215	1.191	0.994	0.650	0.668	1.095	2.205	7.926	1.232
	80%	1.518	0.749	3.820	8.757	2.130	2.014	1.830	1.614	0.878	0.934	0.711	0.932	1.217	1.195	1.791	2.518	2.412
	Sum$_{30,50,80\%}$	1.105	670.6	0.570	1.029	1.476	351.0	2.25	1.23	2.019	1.188	246.8	202.4	1.686	1.215	2.894	26.6	342.9

Muscles are: vastii (V), semitendinosus (SEMI_T), semimembranosus (SEMI_M), sartorius (SART), rectus femoris (RF), gracilis (GRAC), biceps femoris long head (BFL), biceps femoris short head (BFS), adductor magnus (AM), adductor longus (AL), flexor digitorum longus (FDL), gastrocnemius lateralis (GLAT), gastrocnemius medialis (GMED), peroneal group (PER; peroneus longus, brevis and tertius), soleus with flexor hallucis longus (SOL), anterior tibial muscles (ANT; tibialis anterior, extensor digitorum longus, extensor hallucis longus), tibialis posterior (TIBP)

Supplementary Table 5: Results of algorithm 2 testing

Number of Slices	Muscle																
	V	SEMI_T	SEMI_M	SART	RF	GRAC	BFL	BFS	AM	AL	FDL	GLAT	GMED	PER	SOL	ANT	TIBP
For criterion 1: Difference between percentage change in muscle size calculated by sub-algorithm versus percentage change in muscle volume (%)																	
1	-0.190	-0.839	1.086	0.812	0.482	-0.446	-1.741	0.246	0.100	-0.443	-0.402	0.003	0.521	-2.911	-0.532	-0.714	0.036
2	2.242	-0.923	0.726	-0.148	0.720	-0.946	-1.761	0.938	0.344	-0.488	0.629	0.596	0.287	-2.894	-0.841	0.026	1.309
3	2.197	-1.109	0.664	0.005	0.604	-0.571	-1.251	0.853	0.166	-0.195	0.352	0.479	0.069	-2.342	-0.576	0.126	1.096
4	1.874	-1.164	0.855	0.079	0.678	-0.014	-1.340	1.050	0.295	-0.171	0.318	0.643	0.123	-1.908	-0.382	0.343	1.207
5	1.505	-1.049	1.048	0.281	0.568	0.462	-1.177	1.063	0.210	-0.252	0.335	0.547	-0.039	-1.591	-0.287	0.332	1.005
6	1.258	-0.945	1.176	0.211	0.633	0.876	-1.139	1.217	0.223	-0.324	0.344	0.511	-0.052	-1.453	-0.198	0.470	1.067
7	1.037	-0.867	1.356	0.019	0.515	1.235	-1.028	1.232	0.147	-0.405	0.345	0.429	-0.073	-1.279	-0.135	0.489	1.006
8	0.871	-0.751	1.500	-0.148	0.521	1.340	-1.008	1.329	0.192	-0.423	0.406	0.482	-0.013	-1.185	-0.016	0.559	1.051
9	0.718	-0.689	1.683	-0.307	0.390	1.511	-0.945	1.296	0.208	-0.480	0.362	0.445	0.006	-1.015	0.097	0.559	0.906
10	0.609	-0.562	1.759	-0.430	0.346	1.499	-0.880	1.354	0.278	-0.508	0.369	0.393	0.072	-0.885	0.225	0.584	0.836
11	0.498	-0.495	1.885	-0.536	0.242	1.579	-0.802	1.302	0.326	-0.439	0.305	0.477	0.076	-0.720	0.249	0.555	0.623
12	0.418	-0.376	1.930	-0.630	0.212	1.513	-0.753	1.307	0.379	-0.383	0.259	0.477	0.124	-0.615	0.286	0.543	0.477
13	0.343	-0.292	1.976	-0.723	0.055	1.594	-0.684	1.233	0.367	-0.338	0.167	0.466	0.114	-0.433	0.275	0.507	0.299
14	0.287	-0.176	1.914	-0.753	-0.026	1.552	-0.585	1.200	0.389	-0.291	0.127	0.419	0.124	-0.283	0.284	0.493	0.183
15	0.238	-0.097	1.877	-0.763	-0.204	1.649	-0.456	1.117	0.366	-0.256	0.072	0.365	0.092	-0.118	0.256	0.447	0.020
17	0.181	0.061	1.647	-0.699	-0.420	1.704	-0.246	0.991	0.343	-0.197	0.009	0.292	0.051	0.093	0.213	0.360	-0.202
19	0.148	0.166	1.351	-0.612	-0.502	1.691	-0.149	0.854	0.290	-0.156	-0.016	0.220	0.026	0.246	0.157	0.269	-0.332
21	0.133	0.212	1.001	-0.517	-0.463	1.565	-0.112	0.698	0.241	-0.123	-0.035	0.152	0.017	0.376	0.109	0.188	-0.356
23	0.115	0.215	0.737	-0.417	-0.376	1.324	-0.089	0.551	0.193	-0.095							
For criterion 2: Variability (expressed as a fraction of the variability of the percentage change in muscle volume [=1])																	
1	0.882	2.443	1.479	1.610	0.902	1.470	1.095	1.117	0.853	1.167	2.512	2.136	1.002	1.408	1.407	1.030	1.008
2	1.653	2.014	1.297	1.133	0.574	1.323	0.780	0.784	0.733	1.121	2.281	2.212	0.772	1.331	1.131	0.771	1.033
3	1.672	1.819	1.294	1.005	0.451	1.230	0.702	0.686	0.588	0.883	1.964	1.885	0.641	1.014	1.011	0.590	0.799
4	1.389	1.631	1.010	0.920	0.340	0.949	0.555	0.550	0.454	0.905	1.886	1.668	0.517	0.894	0.730	0.575	0.873
5	1.087	1.383	0.908	0.798	0.209	0.764	0.448	0.472	0.333	0.780	1.525	1.287	0.429	0.631	0.641	0.404	0.662
6	0.848	1.161	0.664	0.714	0.147	0.621	0.336	0.379	0.239	0.766	1.334	1.032	0.371	0.514	0.490	0.378	0.629
7	0.645	0.929	0.565	0.453	0.000	0.488	0.287	0.308	0.163	0.626	0.967	0.752	0.282	0.311	0.420	0.198	0.401
8	0.474	0.732	0.356	0.292	0.098	0.421	0.195	0.231	0.084	0.529	0.806	0.580	0.205	0.199	0.334	0.180	0.319
9	0.330	0.541	0.269	0.181	0.210	0.284	0.158	0.150	0.038	0.377	0.516	0.401	0.124	0.000	0.233	0.000	0.103
10	0.205	0.383	0.105	0.117	0.292	0.205	0.063	0.083	0.088	0.260	0.376	0.270	0.060	0.140	0.143	0.067	0.000
11	0.098	0.231	0.059	0.063	0.427	0.079	0.046	0.013	0.116	0.125	0.141	0.138	0.074	0.338	0.000	0.213	0.222
12	0.000	0.132	0.158	0.041	0.520	0.046	0.129	0.073	0.177	0.000	0.000	0.046	0.115	0.482	0.120	0.284	0.363
13	0.086	0.000	0.228	0.062	0.639	0.124	0.193	0.131	0.232	0.114	0.178	0.145	0.166	0.678	0.238	0.440	0.562
14	0.165	0.107	0.311	0.103	0.726	0.184	0.277	0.188	0.308	0.222	0.294	0.256	0.227	0.822	0.337	0.523	0.710
15	0.234	0.167	0.370	0.157	0.823	0.276	0.346	0.246	0.375	0.313	0.426	0.348	0.285	0.985	0.446	0.666	0.869
17	0.359	0.294	0.461	0.288	0.921	0.415	0.456	0.349	0.508	0.473	0.606	0.496	0.419	1.218	0.612	0.875	1.098
19	0.468	0.391	0.527	0.411	0.946	0.543	0.548	0.446	0.616	0.595	0.721	0.617	0.554	1.365	0.734	1.044	1.214
21	0.568	0.480	0.604	0.520	0.940	0.652	0.638	0.541	0.701	0.688	0.788	0.721	0.671	1.414	0.821	1.142	1.223
23	0.663	0.572	0.686	0.615	0.932	0.739	0.720	0.632	0.769	0.761							

See Table 4 for abbreviations

Supplementary Table 6: Results of algorithm 3 testing

| Number of Slices									Muscle								
	V	SEMI_T	SEMI_M	SART	RF	GRAC	BFL	BFS	AM	AL	FDL	GLAT	GMED	PER	SOL	ANT	TIBP
For criterion 1: Difference between percentage change in muscle size calculated by sub-algorithm versus percentage change in muscle volume (%)																	
1	0.091	-0.889	0.441	0.358	0.499	-0.311	-1.506	0.299	-0.057	0.310	-0.068	-0.261	0.465	-2.886	-0.889	-0.564	0.012
2	0.085	-0.915	1.424	0.179	1.015	1.660	-2.101	1.212	0.670	0.159	0.286	0.415	0.787	-1.851	-0.459	0.324	0.892
3	-0.051	-0.591	1.655	0.524	0.627	2.305	-1.585	1.179	0.347	-0.076	0.478	0.119	0.223	-1.359	-0.411	0.253	0.512
4	-0.074	-0.307	1.945	0.310	0.814	2.037	-1.482	1.489	0.587	0.020	0.778	0.428	0.354	-1.221	-0.084	0.489	0.727
5	-0.138	-0.243	2.184	0.067	0.409	2.181	-1.252	1.328	0.597	-0.090	0.631	0.333	0.267	-0.831	0.114	0.449	0.281
6	-0.132	0.008	2.190	-0.152	0.395	1.828	-1.098	1.410	0.680	-0.118	0.569	0.142	0.452	-0.680	0.284	0.480	0.100
7	-0.154	0.137	2.133	-0.347	-0.085	1.991	-0.909	1.194	0.547	-0.133	0.338	0.129	0.350	-0.266	0.234	0.395	-0.249
8	-0.121	0.347	1.989	-0.349	-0.232	1.781	-0.605	1.116	0.511	-0.135	0.255	0.120	0.316	-0.103	0.239	0.365	-0.378
9	-0.095	0.420	1.709	-0.298	-0.423	1.865	-0.425	0.954	0.434	-0.133	0.159	0.099	0.226	0.086	0.181	0.253	-0.553
10	-0.068	0.483	1.298	-0.212	-0.431	1.643	-0.335	0.828	0.355	-0.131	0.090	0.086	0.165	0.143	0.158	0.193	-0.544
11	-0.028	0.445	1.003	-0.140	-0.357	1.463	-0.278	0.666	0.297	-0.128	0.038	0.088	0.120	0.313	0.130	0.098	-0.481
12	-0.032	0.381	0.795	-0.107	-0.281	1.159	-0.230	0.521	0.242	-0.125	0.014	0.091	0.085	0.413	0.115	0.075	-0.360
13	-0.018	0.301	0.633	-0.047	-0.224	0.896	-0.190	0.399	0.199	-0.123	-0.005	0.093	0.057	0.304	0.111	0.042	-0.270
14	-0.036	0.237	0.504	-0.002	-0.178	0.663	-0.156	0.302	0.157	-0.120	-0.031	0.095	0.034	0.203	0.100	0.013	-0.196
15	-0.033	0.178	0.398	0.009	-0.139	0.481	-0.127	0.222	0.121	-0.118	-0.052	0.097	0.016	0.126	0.090	-0.007	-0.136
17	-0.033	0.125	0.310	0.029	-0.106	0.338	-0.102	0.155	0.091	-0.116							
19	-0.031	0.080	0.236	0.014	-0.078	0.219	-0.080	0.098	0.066	-0.114							
For criterion 2: Variability (expressed as a fraction of the variability of the percentage change in muscle volume [=1])																	
1	0.974	4.564	4.858	2.630	4.409	2.927	2.807	1.873	3.119	4.684	10.407	5.635	1.794	3.489	5.504	4.369	4.559
2	1.043	3.940	2.874	1.959	4.688	4.195	2.568	1.484	3.695	4.732	9.577	4.634	1.253	3.567	4.545	6.697	6.012
3	0.901	2.920	2.819	1.707	3.985	3.808	1.891	1.215	3.013	3.840	6.417	2.948	1.094	2.597	4.079	4.833	4.605
4	0.766	2.598	3.008	1.458	4.605	3.617	2.116	1.113	3.218	3.398	6.250	2.593	1.047	2.434	4.439	5.672	4.495
5	0.631	1.851	2.972	1.270	3.813	2.824	1.810	0.845	2.890	2.442	4.224	1.947	0.856	1.683	3.500	3.994	3.337
6	0.459	1.901	3.066	1.118	3.870	2.609	1.711	0.774	2.975	1.861	3.695	1.509	0.698	1.369	3.478	3.838	2.875
7	0.330	1.486	2.761	0.939	2.894	2.074	1.388	0.504	1.928	1.246	2.636	0.973	0.488	0.832	2.241	2.544	2.016
8	0.142	1.502	2.304	0.657	2.689	1.838	0.998	0.396	1.340	0.693	1.892	0.649	0.335	0.718	1.790	2.056	1.681
9	0.030	1.121	1.719	0.464	2.106	1.405	0.640	0.162	0.734	0.297	1.227	0.274	0.000	0.759	1.056	1.252	1.574
10	0.243	0.948	1.117	0.213	1.704	1.118	0.289	0.000	0.316	0.001	0.727	0.000	0.247	0.790	0.658	0.992	1.471
11	0.372	0.563	0.650	0.063	1.237	0.676	0.000	0.198	0.000	0.232	0.358	0.210	0.444	0.601	0.279	0.703	0.909
12	0.550	0.286	0.286	0.270	0.770	0.316	0.225	0.376	0.244	0.417	0.001	0.376	0.599	0.429	0.002	0.474	0.450
13	0.655	0.000	0.000	0.470	0.342	0.000	0.406	0.525	0.441	0.569	0.289	0.511	0.724	0.243	0.234	0.099	0.001
14	0.733	0.245	0.232	0.723	0.001	0.258	0.554	0.644	0.607	0.696	0.517	0.622	0.828	0.391	0.425	0.347	0.350
15	0.808	0.452	0.423	0.860	0.280	0.464	0.678	0.742	0.749	0.803	0.714	0.715	0.915	0.625	0.586	0.628	0.636
17	0.887	0.627	0.583	0.938	0.514	0.632	0.782	0.824	0.870	0.896							
19	0.956	0.775	0.720	0.940	0.711	0.773	0.872	0.894	0.974	0.975							

See Table 4 for abbreviations

Supplementary Table 7: Results of algorithm 5 testing

Average of every n^{th} slice	Muscle																
	V	SEMI_T	SEMI_M	SART	RF	GRAC	BFL	BFS	AM	AL	FDL	GLAT	GMED	PER	SOL	ANT	TIBP
For criterion 1: Difference between percentage change in muscle size calculated by sub-algorithm versus percentage change in muscle volume (%)																	
2	0.402	-0.89	-0.187	0.534	-0.564	0.704	0.466	0.605	2.524	0.072	0.602	0.602	-1.578	-1.004	0.485	-0.768	-0.796
3	0.522	0.996	-1.03	0.793	0.574	-0.291	1.891	-0.369	1.897	-0.105	0.877	0.877	-0.952	-0.94	-0.137	-1.422	-0.407
4	1.3	-1.695	0.038	1.065	0.475	2.468	0.446	0.289	1.73	-1.311	0.445	0.445	-1.386	-1.733	0.971	-1.2	0.275
5	-2.086	-1.195	-2.771	0.328	0.97	4.087	0.008	2.176	-0.418	1.237	0.216	0.216	0.296	-1.132	-1.504	0.084	2.92
6	0.25	0.623	-0.77	1.323	-0.838	-0.105	2.634	0.158	6.312	-2.496	0.698	0.698	-2.598	-1.235	1.664	-1.221	-0.935
7	1.058	-0.275	-0.772	-0.061	1.821	4.072	1.05	-1.299	4.426	2.938	-0.406	-0.406	-0.672	-0.508	-1.372	-0.327	0.509
8	-0.435	1.716	1.541	0.18	0.133	1.493	-1.472	0.747	6.307	0.773	2.056	2.056	-3.021	-0.649	0.302	-1.958	-4.185
9	1.051	1.171	3.078	-0.165	1.173	1.218	5.188	-2.277	9.741	3.851	0.399	0.399	-4.764	0.153	-3.11	0.11	-0.942
10	0.358	-3.964	0.528	1.875	3.12	11.9	1.684	1.787	3.855	2.635	-0.68	-0.68	-2.484	0.72	-2.609	0.325	6.824
11	-0.88	-2.475	-0.416	0.248	0.946	8.773	-0.949	-1.352	3.898	3.441	-1.578	-1.578	-1.167	3.986	0.349	0.914	-3.833
12	0.086	-0.354	-2.313	0.648	0.24	5.819	-0.116	-4.447	3.658	-1.957	1.097	1.097	0.148	1.37	1.873	-0.919	-2.805
13	-0.461	0.738	-3.535	-1.545	-0.01	4.965	-0.791	-6.955	3.072	-4.272	2.941	2.941	0.783	2.568	-3.511	0.725	-1.877
14	1.508	-2.204	-9.218	1.044	0.539	4.865	-3.139	-4.883	7.92	15.459	3.192	3.192	1.845	1.556	-3.998	1.5	-2.829
15	-0.109	0.944	2.093	4.144	-0.318	3.202	-5.831	0.5	3.35	19.317	4.843	4.843	-1.299	1.857	-4.643	2.267	-4.365
16	0.618	1.285	0.054	0.434	-0.17	1.382	-5.041	-3.005	13.32	19.909	4.575	4.575	-7.331	1.582	-4.556	2.084	-4.429
For criterion 2: Variability (expressed as a fraction of the variability of the percentage change in muscle volume [=1])																	
2	1.417	1.133	1.270	0.754	1.129	1.115	1.846	1.990	1.471	0.681	0.853	1.285	1.498	0.910	1.041	0.976	0.672
3	0.973	2.314	2.250	1.319	1.081	1.026	4.150	3.027	2.867	1.747	2.005	2.265	1.996	1.847	1.475	1.280	1.236
4	1.364	2.524	3.168	1.566	1.657	1.469	3.473	4.256	2.857	1.879	1.699	2.080	2.772	1.673	1.195	1.233	1.267
5	2.784	3.385	4.183	2.242	1.989	2.227	4.226	5.906	3.758	2.326	2.446	4.003	2.898	2.731	1.244	2.484	2.157
6	2.073	3.620	3.355	2.681	1.315	1.121	9.651	3.204	4.988	2.919	2.518	3.911	3.020	3.032	1.921	2.321	1.310
7	2.095	4.909	4.695	2.465	2.177	2.355	4.756	4.498	4.895	2.857	2.488	6.367	5.096	2.431	2.792	1.631	2.524
8	1.936	3.135	6.422	2.309	1.215	1.397	5.609	7.702	6.232	2.833	3.138	3.351	5.227	3.778	2.385	3.138	1.420
9	3.011	6.572	6.138	2.380	1.802	2.356	12.863	5.146	7.305	3.938	3.768	3.026	4.002	4.102	1.518	1.245	1.742
10	2.632	6.471	4.797	5.044	4.196	5.216	10.225	6.447	3.758	6.644	2.461	2.872	2.337	4.495	2.326	2.005	4.877
11	3.209	5.479	6.270	2.387	3.625	3.885	11.811	5.220	3.854	12.735	3.148	6.655	3.600	7.434	5.300	3.798	4.509
12	4.104	5.978	5.267	2.173	1.652	2.485	9.475	7.226	4.222	12.386	3.410	9.962	5.292	8.135	4.073	4.215	3.139
13	7.812	5.519	6.507	3.076	1.317	2.719	9.105	11.187	6.014	11.474	4.383	8.552	5.760	7.445	2.620	3.102	3.380
14	4.500	6.133	11.216	4.893	1.618	2.660	11.227	14.652	12.742	29.558	4.446	23.133	13.530	6.329	3.347	3.270	3.479
15	2.019	7.261	30.085	7.815	2.743	2.751	16.320	20.725	12.816	33.098	10.764	25.318	13.522	6.744	3.709	5.159	3.049
16	1.631	6.998	31.938	4.524	2.979	3.864	21.445	22.706	16.154	33.051	11.045	16.915	13.309	6.526	4.271	5.705	3.247

See Table 4 for abbreviations
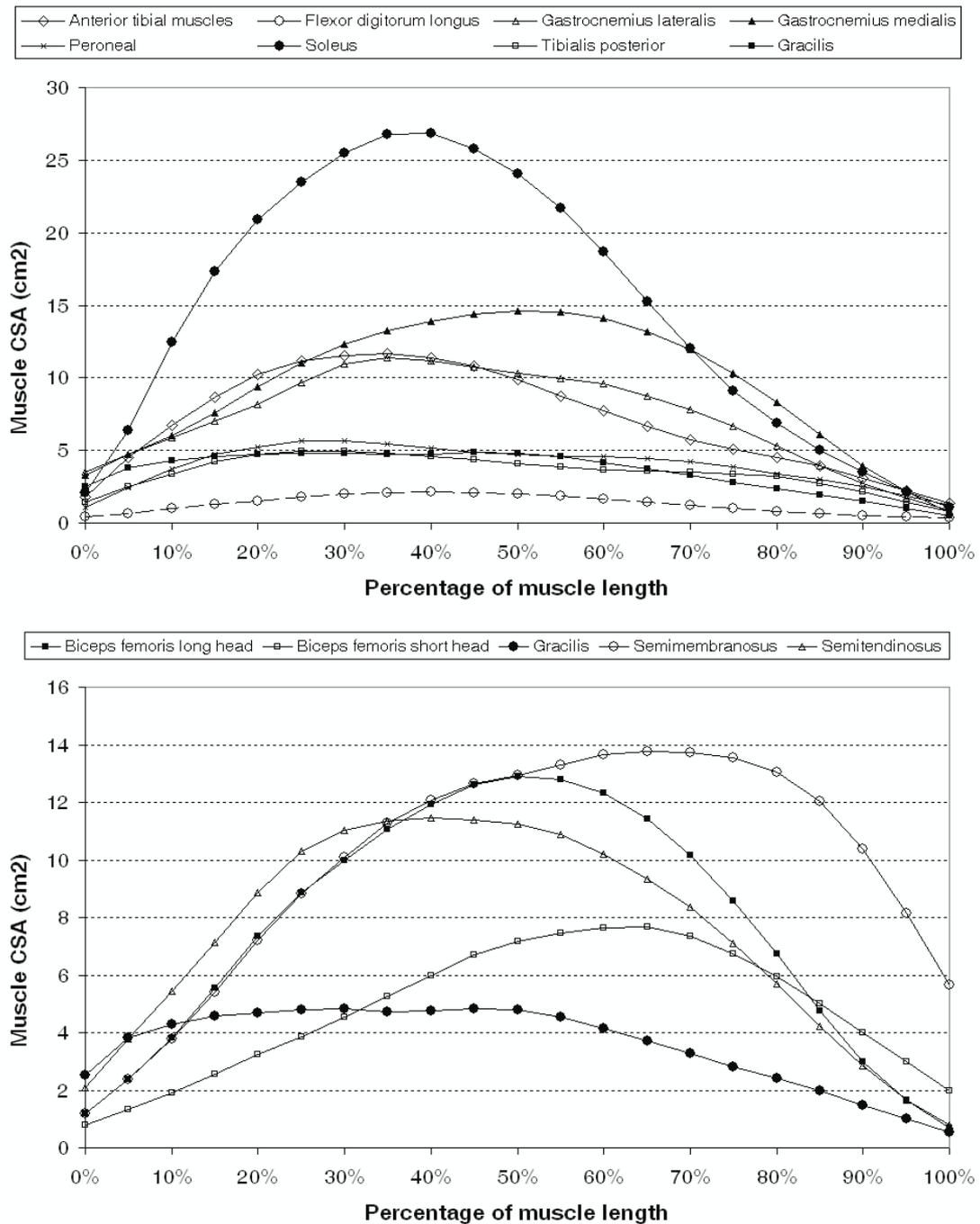
Supplementary Figure 4: Profiles of the calf (top) and five thigh (bottom) muscles. Data were resampled from original cross-sectional area (CSA) data such that 19 measurements were interpolated between the first (0%) and last (100%) CSA measurements. Average CSA profiles were then generated for each subject from all his measurements and then averaged between subjects to generate the data presented here.
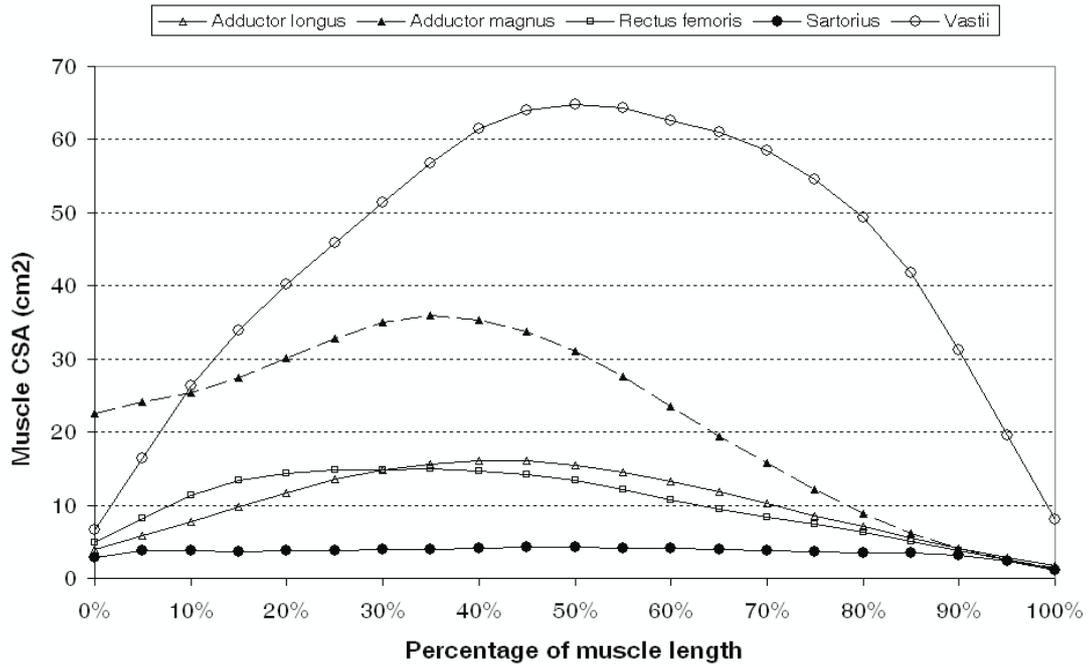
Supplementary Figure 5: Profiles of the five remaining thigh muscles. Data were resampled from original cross-sectional area (CSA) data such that 19 measurements were interpolated between the first (0%) and last (100%) CSA measurements. Average CSA profiles were then generated for each subject from all his measurements and then averaged between subjects to generate the data presented here. Note that it is likely a (proximal) portion of adductor magnus was not captured in the imaging process.