Reference Values for Volume, Fat Content and Shape of the Hip Abductor Muscles in Healthy Individuals from Dixon MRI

Martin A. Belzunce, PhD[1*], Johann Henckel, MD[1], Anna Di Laura, PhD[1], and Alister J. Hart, MD[1,2]

1 Royal National Orthopaedic Hospital, Stanmore HA7 4LP, UK

2 Institute of Orthopaedics and Musculoskeletal Science, University College London, Stanmore HA7 4LP, UK

*Corresponding author: Martin A. Belzunce at Research, Royal National Orthopaedic Hospital (RNOH), Brockley Hill, Stanmore, Middlesex, HA7 4LP. E-mail: martin.belzunce@nhs.net.

Keywords: magnetic resonance imaging, hip muscles, intramuscular fat, atlas, muscle volume, fat fraction, Dixon

List of Abbreviations

CSA: cross-sectional area, FOV: field of view, GMAX: gluteus maximus, GMED: gluteus medius, GMIN: gluteus minimus, IMF: intramuscular fat, IQR: interquartile range, LBM: lean body mass, LMV: lean muscle volume, NV: normalized volume, SD: standard deviation, TE: echo time, TFL: tensor fascia latae, TR: repetition time

Abstract

Healthy hip abductor muscles are a good indicator of a healthy hip and an active lifestyle as they are greatly involved in human daily activities. Fatty infiltration and muscle atrophy are associated with loss of strength, mobility and hip disease. However, these variables have not been widely studied in this muscle group. We aimed to characterise the hip abductor muscles in a group of healthy individuals to establish reference values for volume, intramuscular fat content and shape of this muscle group. To achieve this, we executed a cross-sectional study using Dixon MRI scans of 51 healthy subjects. We used an automated segmentation method to label GMAX, GMED, GMIN and TFL muscles, and measured normalized volume (NV) using lean body mass, fat fraction (FF) and lean muscle volume (LMV) for each subject and computed non-parametric statistics for each variable grouped by sex and age. We measured these variables for each axial slice and created cross-sectional area (CSA) and FF axial profiles for each muscle. Finally, we generated sex specific atlases with FF statistical images. We measured median (IQR) NV values of 12.6 (10.8-13.8), 6.3 (5.6 – 6.7), 1.6 (1.4-1.7) and 0.8 (0.6-1.0) cm^3/Kg for GMAX, GMED, GMIN and TFL; and median (IQR) FF values of 12.3 (10.1-15.9)%, 9.8 (8.6-11.2)%, 10.0 (9.0-12.0)% and 10.2 (7.8-13.5)% respectively. FF values were significantly higher for females for the four muscles ($p<0.01$), but there were no significant differences between two age groups. When comparing individual muscles, we observed a significant higher FF in GMAX than in the other muscles. The reported novel reference values and axial profiles for volume and FF of the hip abductors, together with male and female atlases, are tools that could potentially help to quantify and early detect the deteriorating effects of hip disease or sarcopenia.

Introduction

Healthy hip abductor muscles are a good indicator of a healthy hip and an active lifestyle as they are fundamental to human daily activities, such as standing, walking and running [1,2]. Fatty infiltration and muscle wasting (atrophy) are associated with loss of strength and mobility, making volume and intramuscular fat (IMF) content important markers for muscle health [3–13].

Magnetic resonance imaging (MRI) is the best imaging modality to assess these variables both qualitative and quantitatively, in particular Dixon MRI that provides a fat fraction (FF) signal that can be used to quantify IMF infiltration. In clinical practice, fatty atrophy is usually classified with the Goutallier grading scale [14] and muscle waste by visual inspection or measuring cross-sectional areas (CSA) [15,16]. However, the introduction of new automated methods for segmenting and labelling the hip muscles from MR images [17–21] has opened the possibility of performing full 3D quantitative muscle assessment from MRI scans without the time-demanding and impractical manual labelling.

Characterising the hip abductor muscles in a group of healthy individuals can provide new reference values of volume and fat infiltration for the healthy hip abductors, which could help to quantify and early detect the deteriorating effects of hip disease or sarcopenia. However, reference data for the hip abductors in healthy individuals is limited and incomplete. Volume and fat fraction reference values of healthy subjects have been only reported for gluteus medius (GMED) and minimus (GMIN) [22]. Tensor fascia latae (TFL) and maximus (GMAX) also contribute to hip abduction, in addition to other important functions during the gait cycle, but reference data have not been published for them. These four muscles have previously studied in diseased hips [14,23], in elderly people [24,25] and in patients that underwent hip arthroplasty [26], but they have been mainly assessed by visual inspection for fat infiltration and measuring CSAs for muscle size.

In this work, we aimed to characterise the main abductor muscles from 51 healthy individuals that underwent Dixon MRI. We focus on GMED, GMIN and TFL that are essential to stabilize and control the pelvis during the gait cycle, and we also include GMAX that assists in hip abduction in addition to its main function of extending the hip [2]. To do this, we not only report reference values for volume and IMF content, analysing the effects of sex and age; but also provide shape measurements, axial profiles of CSAs and FF for each muscle, including male and female atlases that models the shape and fat content from our study sample.

Materials and methods

Study Design

This was a cross-sectional study looking at the effects of marathon running in the hip muscles and joints (approved by the Institutional Research Ethics Committee [13823/001]). The study was conducted in London using MRI scans from 51 healthy subjects. The inclusion criteria for the participants of this study was the absence of hip injury or hip surgery and no contraindication to MRI. Of the 51 subjects, 8 were non runners (doing no more than 90 minutes of physical activity per week), 27 were in the first weeks of a training programme for their first marathon (running recreationally at least 2 times a week at the time of the scan) and 16 were experienced runners that had run at least three marathons or ultra-marathons in the past. Indistinctly of their level of physical activity, the three groups had similar mean (SD) BMI values: 22.0 (2.6), 23.2 (2.1) kg/m^2 and 23.8 (3.0) kg/m^2 respectively. The demographic characteristics of the study sample are shown in Table 1. All subjects provided written informed consent. All procedures performed in this work involving human participants were approved by the local Institutional Review Board.

The MRI protocol consisted of standard clinical sequences for the hips and axial Dixon with a field of view (FOV) that covered axially from the lesser trochanter to the top of the iliac crest. In this study, we only use the Dixon images because they provide fat and water images which can be used to estimate a quantitative FF signal [27–30]. The MRI scans were acquired on a 3T scanner (Siemens Magneton Vida, Erlangen, Germany) using a body coil. The Siemens commercial two-points TSE-Dixon sequence[31] was used with the following parameters: slice thickness 1.5 mm, spacing between slices 1.95 mm, repetition time (TR) 4570 msec, echo time (TE) 45 msec, number of excitations 1, number of echoes 14, flip angle 120°. The voxel size was $0.47 \times 0.47 \times 1.95\ mm^3$.

Volume, Fat Content and Shape Measurements

Volume, fat content and shape were estimated from 3D labels of GMAX, GMED, GMIN and TFL muscles for each subject, which were created using an automated method that labels left and right hip abductor muscles in the in-phase Dixon image. The automated segmentation tool is an in-house plugin for Simpleware™ ScanIP (Version 2020.6; Synopsys, Inc., Mountain View, USA), which is based on a multi-atlas segmentation method that employs a library with 15 manually segmented Dixon scans [17,18]. The automated labels were verified by an experienced user and manually corrected when necessary. All the measurements were assessed globally, by sex and by age. For the latter, we divided the study sample into two age groups, 18-39 and 39-60 years with N=38 (24 males, 14 females) and N=13 (7 males, 6 females) respectively.

Volume

The volume was estimated as the sum of the voxels in each label. As muscle volume is generally proportional to body mass, we also provide normalized volume (NV) values where the volume measures were divided by the lean body mass (LBM) of each subject. The latter was estimated using the Boer formula [32]:

$$LBM_M = -.407W + 0.267H - 19.2$$

$$LBM_W = -.252W + 0.473H - 48.3$$

where LBM_M and LBM_W are the lean body mass in Kg for males and females respectively, W is the weight in Kg and H the height in cm for each subject.

In addition, cross-sectional areas (CSA) for each axial slice that forms a label were computed to obtain size profiles for each muscle. CSAs were also normalized by LBM for the profiles.

Since the number of slices covered by each muscle varied for each subject, we resampled all the profiles (with a different number of slices for each subject) into 50 fixed slices or sampling points by applying a linear interpolation. In the resampled profiles, the first and most superior slice is the origin of the muscle and slice number 50 is the insertion and most inferior slice. The mean number of slices in the original profiles (before resampling) were 107, 89, 60 and 71 for GMAX, GMED, GMIN and TFL respectively, resulting in mean resampled slices width of 4.3, 3.5, 2.4 and 2.8 mm (from the original 1.95 mm).

We computed the median, the interquartile range (IQR) and the central 90% interval for each slice of the resampled CSAs profiles. The 5 axial slices with the highest mean CSA values were identified for each muscle as this is useful information for studies using CSAs to assess muscle [33]. To generate smoother profiles, a moving average filter with a window of three slices was applied to each of them.

Fat Fraction Measurement

The muscles labels from the in-phase Dixon image were transferred to the FF image, which is the ratio between the fat image and the sum of the water and fat images. The mean FF signal in each label was computed as a measure of fat content in each muscle as described in [17]. This metric quantifies IMF. The FF values were computed also for each axial slice following the same slice resampling as for the CSAs to generate FF profiles. Median, interquartile range (IQR) and the central 90% interval values were estimated for each slice.

Lean Muscle Volume Measurement

The lean muscle volume (LMV) was computed for each subject using the previously computed volume and FF metrics:

$$LMV_{im} = V_{im}(1 - FF_{im})$$

where LMV_{im} is the lean muscle volume of muscle m for subject i, and V_{im} and FF_{im} the respective full volume and fat fraction values. LMV values were also normalized by LBM and normalized LMV profiles were computed in the same fashion as volume and FF profiles.

Muscle Shape

The previously described CSA profiles provide information about muscle shape along the axial axis. In addition, we measured the shape factor for each muscle, which is defined as the ratio between the mean CSA and the maximum CSA. This metric informs if the maximum CSA multiplied by the muscle length can be used as a surrogate metric for muscle volume [34,35].

As a measure of muscle shape, we fitted an ellipse to each 2D muscle label at the level of the maximum CSA. Each ellipse had the same normalized second central moments as the region defined by each muscle label. The minor and major axes of each ellipse, as well as their orientation (measured as the angle between the major axis and the x axis of the image), were measured. Finally, a bounding box was computed for each 3D label to inform the profiles about the muscle extension and proportion in each direction: width (x), depth (y) and height (z).

Atlas for Fat Fraction Distribution

We generated male and female atlases of the pelvis to characterize the IMF content of the hip abductor muscles employing the 31 and 20 available Dixon scans respectively. Each Dixon images were segmented into air, soft-tissue, mixed soft-tissue/fat and fat/bone tissue as described in [36] and labels for GMAX, GMED, GMIN and TFL were added.

For the female atlas, the Dixon MRI of a female with a BMI of 22.2 kg/m2, height of 1.67 m and average muscle volumes (between left and right) of 584 cm^3, 281 cm^3, 76 cm^3 and 52 cm^3 was selected as a reference. The female segmented images were registered to the reference image with a B-spline non-rigid registration [37] using SimpleElastix [38,39]. A majority voting algorithm was used to fuse the registered labels into the final atlas labels. Using the same non-

rigid transform, the fat fraction images of each subject were propagated into the atlas space and median and percentile (10^{th}, 25^{th}, 75^{th} and 90^{th}) fat fraction images were estimated for each muscle. The same process was repeated for the male group.

In addition, we propose a processing chain, described in Figure 1, which uses the gender-specific atlas to detect regions of high IMF content for any MRI scan used as input. In this processing scheme, the input Dixon MRI is segmented and then registered to the atlas as described above for the atlas generation. Then the FF image, generated from the Dixon input, is propagated into the atlas space and compared to the median and percentile FF images of the atlas. As a result, an image with the FF difference between the input and the median of the atlas is obtained, together with masks for the regions where the input FF was higher than the 75^{th} and 90^{th} percentile FF image of the atlas. We applied this method to 5 scans.

Statistical Analyses

Volume, NV, FF and LMV values were tested for normality within each group using a Shapiro-Wilk test. For some of the groups, the null hypothesis that the group's samples followed a normal distribution was rejected. For this reason, we reported parametric (mean ± standard deviation (SD)) and non-parametric (median and interquartile range (IQR)) descriptive statistics for these variables.

The effect of gender and age on muscle size and fat content was assessed using a Kruskal-Wallis test for each muscle. For these tests, the NV values were used. We used eta squared (η^2) for Kruskal-Wallis as a measure of effect sizes[40], with Cohen[41] reference values to define small ($\eta^2 = 0.01$), medium ($\eta^2 = 0.06$), and large ($\eta^2 = 0.14$) effects.

Outliers were defined as individual fat fraction or NV values that were more than three scaled median absolute deviations (MAD) away from the median. Subjects with more than half of the muscles with outlier values were identified.

We used a level of statistical significance (α) of 0.05 for all the tests.

Results

Reference IQR intervals for volume, NV and FF are shown in Table 2. We found that the muscle volume of GMAX, GMED, GMIN and TFL were higher in males than in females, but when the volume was normalized by LBM there was not any significant difference for GMAX and TFL between sex. FF was significantly higher in females than males for all four muscles, but there was no significant difference between the two age groups. When comparing individual muscles, we observed a significant higher FF in GMAX than in the other muscles.

Volume

The median (IQR) muscle volumes were 650 (555-775) cm^3, 332 (280-406) cm^3, 86 (70-99) cm^3 and 43 (34-59) cm^3 for GMAX, GMED, GMIN and TFL respectively; while the median (IQR) normalized volume values were 12.6 (10.8-13.8) cm^3/Kg, 6.3 (5.6-6.7) cm^3/Kg, 1.6 (1.4-1.7) cm^3/Kg and 0.8 (0.6-1.0) cm^3/Kg. Table 2 contains the full volume results, including descriptive statistics for each sex.

The volume values did not follow a normal distribution for any of the muscles ($p=0.03$, $p=0.03$, $p<0.01$ and $p<0.01$ for GMAX, GMED, GMIN and TFL respectively), while the same was observed for the normalized volume values only for GMIN ($p<0.01$) and TFL ($p<0.01$). The NV distribution of the study sample can be seen in the supplementary material (Figure S.1).

Volume was higher correlated to LBM ($r=0.62$, $r=0.81$, $r=0.85$ and $r=0.58$ for GMAX, GMED, GMIN and TFL respectively) than to height ($r=0.50$, $r=0.67$, $r=0.75$ and $r=0.39$).

The NV values are presented in Figure 2 with boxplots for sex and age groups. Males had a statistically significant higher NV than females for GMED ($p<0.01$, medium effect size, $\eta^2=0.11$), GMIN ($p=0.02$, small effect size, $\eta^2=0.04$) and TFL ($p<0.01$, small effect size,

η^2=0.03), but no significant differences were found for GMAX (p=0.9). There were no significant differences between the age groups for each muscle (p=0.46, p=0.73, p=0.10 and p=0.20).

Three subjects had at least two muscles with outlier NV values, but none of them had more than four muscles in the outlier region. Two of them had bilateral large TFL muscles and one bilateral large GMIN muscles.

Fat Fraction

The median (IQR) FF values were 12.3 (10.1-15.9) %, 9.8 (8.6-11.2) %, 10.0 (9.0-12.0) % and 10.2 (7.8-13.5) % for GMAX, GMED, GMIN and TFL and are shown in Figure 3 with boxplots. Table 2 show FF mean (SD) and median (IQR) values by sex. The Shapiro-Wilk test for normality was rejected (p<0.01 for the four muscles). The FF values had a right-skewed distribution that can be seen in Figure S.2 in the supplementary material.

FF was significantly higher in females than in males for GMAX (p<0.01, large effect size, η^2=0.17), GMED (p<0.01, medium effect size, η^2=0.08), GMIN (p<0.01, large effect size, η^2=0.19) and TFL (p<0.01, large effect size, η^2=0.28). On contrary, there were no significant differences between the two age groups (p=0.08, p=0.48, p=0.12 and p=0.05 for GMAX, GMED, GMIN and TFL respectively). GMAX FF was significantly higher than in the other muscles (Kruskal-Wallis test, p <0.01, medium effect size, η^2=0.10), but there were no significant differences between GMED, GMIN and TFL.

Six subjects had at least one muscle with outlier FF values. Only one subject had all their muscle FF measurements in the outlier criteria (high fat infiltration) and they were identified as an outlier.

Lean Muscle Volume

The median (IQR) normalized LMV was 12.1 (9.9-13.1) cm^3/Kg for GMAX, 6.0 (5.4-6.5) cm^3/Kg for GMED, 1.5 (1.4-1.6) cm^3/Kg for GMIN and 0.8 (0.6-0.9) cm^3/Kg for TFL. The normalized LMV values were significantly different between males and females for GMED ($p<0.01$), GMIN ($p=0.01$) and TFL ($p<0.01$), but they were not for GMAX ($p=0.89$). These differences are a result of the higher FF in females than in males. There were no significant differences between the two age groups ($p=0.54$, $p=0.95$, $p=0.42$ and $p=0.49$).

Cross-Sectional Areas and Fat Fraction Profiles

Axial profiles of normalized CSA, FF and normalized lean CSA are shown in Figure 4 using the median value and the IQR interval as error bar for each slice, starting at the origin of the muscle (slice 1) and ending in the insertion (slice 50). For GMAX and TFL, the profile ends in the lesser trochanter, before the insertion of the muscle due to the FOV used.

In Figure S.3 of the supplementary material, we have also included axial profiles with error bars for the central 90% interval ranges instead of the IQR. In the FF profiles, higher fat content and variability across individuals is observed for GMAX than for GMED and GMIN.

In the axial profiles, the regions with the highest lean muscle CSAs were identified in red, which show which sections of each muscle are the muscle bulk. These regions are shown for a single subject in Figure 5.

Muscle Shape

The mean (SD) shape factors were 0.64 (0.05), 0.52 (0.05), 0.35 (0.04) and 0.40 (0.05) for GMAX, GMED, GMIN and TFL respectively. For the ellipses fitted at the slice of maximum CSA, the mean (SD) major axis lengths were 166 (14) mm, 114 (15) mm, 85 (11) mm and 39 (7) mm for GMAX, GMED, GMIN and TFL respectively; while the minor axes were 46 (6)

mm, 47 (8) mm, 23 (4) mm and 18 (3) mm long. The orientation angle of the ellipses had mean (SD) values of 16° (5°), 54° (7°), 65° (8°) and -44° (14)°. In Figure S.4 of the supplementary material, all the fitted ellipses are plotted overlaid on the MRI image of a single subject. The mean ellipses for males and females are also shown.

In terms of the extent of each muscle in the image coordinate system, the average bounding box covering each muscle was (154, 91, 187) cm^3, (95, 114, 178) cm^3, (66, 86, 121) cm^3 and (49, 42, 115) cm^3 for GMAX, GMED, GMIN and TFL. GMAX and TFL sizes in *z* are cropped at the lesser trochanter due to the FOV used for the MRI scans.

Atlas for Fat Fraction Distribution

In Figure 6, coronal and axial images of the atlas are shown for the labels image (A) and the median FF image (B). The atlas also contains images for the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles of FF. The labels image and the median FF image for the male atlas are available in the supplementary material (Figure S.5). The full atlas is publicly available at http://doi.org/10.5281/zenodo.4063698.

In Figure 7, we show images with the FF difference between the input and the median of the atlas for 5 cases that were obtained by applying the proposed processing chain. The images show where the increased fat accumulation is spatially located for the two cases with the highest GMAX FF values in our study sample (subjects 9 and 33).

Discussion

This is the first study to quantitatively measure volume and IMF content of the main hip abductor muscles simultaneously (GMAX, GMED, GMIN and TFL) in healthy individuals. In addition, we analysed the spatial distribution of these variables, providing unique and novel information about muscle shape and IMF distribution within each muscle. We report normalized volume and FF (as a measure of IMF content) intervals that can be potentially used as a reference in research and clinical practice. Intervals for the different sexes are needed, as we found that FF was significantly higher in females than males, although we did not find significant sex differences for normalized volume in GMED and GMIN. In addition, axial profiles of normalized CSA and FF are presented for each muscle, which provides a better understanding of how these variables change throughout the muscle. The generated reference atlases for males and females model the average shape and fat content of each muscle, showing the spatially heterogeneous distribution of IMF in healthy individuals. These data, together with automated muscle labelling tools and the wider availability of MRI scans, have the potential to translate into new clinical tools that can assist radiologist with quantitative measurements from MRI of the pelvis.

Normative FF and volume values have been previously reported only for gluteus medius and minimus [22]. Mean FF values of 8 and 10% were found for these two muscles respectively and female subjects had higher FF, but the differences between sexes were not significant. In our study, we measured slightly higher FF values with median values of 10-11%, although we did find significant differences between males and females. Both studies included individuals with healthy BMI values (mean BMI of 22.9 and 23.4 kg/m^2 for males and females in our study, while 22.8 and 21.6 kg/m^2 respectively in [22]), although the mean BMI was lower for female subjects in [22], which could potentially explain the gender differences between the two studies. Our marginally higher FF values could have been the result of using a different implementation

of the Dixon sequence or due to differences in the segmentation. In terms of volume values, we normalized the volume values by the LBM of each subject (while the squared height was used in [22]) because we considered that LBM is a more representative metric of body size. Nevertheless, to draw a comparison with [22], we normalized our data by the squared height of each subject and we obtained similar results for GMED and GMIN (101 and 28 cm^3/m^2 respectively, compared to 105 and 30 cm^3/m^2 in [22]). We also found that GMAX had a higher IMF content than GMED and GMIN. Differences in IMF content of individual muscles within a muscle group have also been previously detected for the calves [5,13].

Having reference intervals for IMF content and volume in healthy individuals in the hip abductors is important since pronounced fatty infiltration in the gluteal muscles have been correlated with the severity of OA [14] and linked to falls in the elderly [24]. We not only provide reference values for each muscle but also their spatial distribution with axial profiles, shape metrics and atlases that provide new ways and data to assess the hip abductors. The atlases can be used to execute voxel-based and regional analysis of the IMF content of a given subject relative to our sample of healthy subjects, as it has been proposed in other applications [42].

The CSA profiles and shape metrics allow the comparison of our results with studies using only 2D CSAs, as the reported shape factors are usually used to extrapolate volume from a single CSA and single CSAs can be compared to their respective slice in the profiles. Furthermore, we proposed the use of a fitted ellipse at the slice of maximum CSA as a simplified model for muscle shape, which captures the size and orientation of each muscle in the standard MRI supine position.

The main limitation of this work is the small sample size of the study, which makes the presented reference values less accurate. Large sample sizes in MR imaging studies are difficult to achieve, especially for specific sequences that are not widely used clinically. Despite the

relatively low sample size, we got significantly different FF values between males and females, with high effect sizes in three out of the four muscles under study. We also observed significant FF differences between GMAX and the other muscles (both tests with $p < 0.01$), with medium effect sizes. The medium and large effect sizes for FF, increase the likelihood that similar results could be found for a larger sample. On contrary, we obtained low effect sizes when comparing NV between groups. Therefore, a higher sample size would be needed to clarify if detecting volume sex differences in only three of the four hip abductors was due to low statistical power in our tests. In addition, we included histograms with normalized volume and FF distributions in the supplementary material that show that volume follows approximately a normal distribution and FF follows a skewed distribution [43].

A second limitation of our study design is that we only included subjects aged under 60. However, this allowed us to report reference values for healthy subjects not affected by sarcopenia [44]. We divided the subjects into two age subgroups and we did not find any significant difference between them, although the size of the groups, when divided by age and sex, were small to be conclusive. Similar finding where reported in [22] for GMED and GMIN for a larger sample size.

A third limitation is that the FOV of our MRI protocol starts at the level of the lesser trochanter, leaving out a small part of the GMAX and TFL muscles and resulting in smaller volume values. However, the use of the lesser trochanter to limit the FOV allows the standardization of the measurements and avoids unnecessarily longer scans to cover the full insertion of GMAX and TFL. In terms of FF, the GMAX and TFL profiles of Figure 4 show that the fat infiltration is not increasing at the level of the lesser trochanter and therefore the cropping of a small part of these muscles would not affect the reported FF values.

Finally, we used an automated method to label the muscles, which are less accurate than manually delineated labels. However, the error of the automated labelling is low for this type of cross-sectional study [17] and, in addition, the labels were manually corrected when suboptimal segmentations were observed. The quality of the segmentations can be observed in Figure S.6 of the supplementary material.

To conclude, we report novel quantitative data of the hip abductors in healthy individuals from Dixon MRI, including reference interval values for normalized volume and FF; axial profiles with reference interval values for normalized CSA, FF and lean CSA; including male and female atlases with median and percentile images. These data sets can be used as a reference for healthy abductors in clinical research and to develop novel tools for a comprehensive and automated assessment of this muscle group.

Acknowledgments

This study was funded by patient donations, the Royal National Orthopaedic Hospital (RNOH) Charity, the Rosetrees & Stoneygate Trusts, the Maurice Hatter Foundation, Trustees of the London Clinic Charity and supported by researchers at the National Institute for Health Research University College London Hospitals Biomedical Research Centre (NIHR UCLH BRC).

Data Availability Statement

The atlases images generated in this study are openly available at http://doi.org/10.5281/zenodo.4063698.

References

1. Anderson FC, Pandy MG. Individual muscle contributions to support in normal walking. *Gait Posture*. 2003;17(2):159-169. doi:10.1016/S0966-6362(02)00073-5
2. Bartlett JL, Sumner B, Ellis RG, Kram R. Activity and functions of the human gluteal muscles in walking, running, sprinting, and climbing. *Am J Phys Anthropol*. 2014;153(1):124-131. doi:10.1002/ajpa.22419
3. Addison O, Marcus RL, Lastayo PC, Ryan AS. Intermuscular fat: A review of the consequences and causes. *Int J Endocrinol*. 2014;2014. doi:10.1155/2014/309570
4. Marcus RL, Addison O, Kidde JP, Dibble LE, Lastayo PC. Skeletal muscle fat infiltration: Impact of age, inactivity, and exercise. *J Nutr Heal Aging*. 2010;14(5):362-366. doi:10.1007/s12603-010-0081-2
5. Commean PK, Tuttle LJ, Hastings MK, Strube MJ, Mueller MJ. Magnetic resonance imaging measurement reproducibility for calf muscle and adipose tissue volume. *J Magn Reson Imaging*. 2011;34(6):1285-1294. doi:10.1002/jmri.22791
6. Hilton TN, Tuttle LJ, Bohnert KL, Mueller MJ, Sinacore DR. Excessive Adipose Tissue Infiltration in Skeletal Muscle in Individuals With Obesity, Diabetes Mellitus, and Peripheral Neuropathy: Association With Performance and Function. *Phys Ther*. 2008;88(11):1336-1344. doi:10.2522/ptj.20080079
7. Goodpaster BH, Chomentowski P, Ward BK, et al. Effects of physical activity on strength and skeletal muscle fat infiltration in older adults: A randomized controlled trial. *J Appl Physiol*. 2008;105(5):1498-1503. doi:10.1152/japplphysiol.90425.2008
8. Kumar D, Karampinos DC, MacLeod TD, et al. Quadriceps intramuscular fat fraction rather than muscle size is associated with knee osteoarthritis. *Osteoarthr Cartil*. 2014;22(2):226-234. doi:10.1016/j.joca.2013.12.005

9. Manini TM, Clark BC, Nalls MA, Goodpaster BH, Ploutz-Snyder LL, Harris TB. Reduced physical activity increases intermuscular adipose tissue in healthy young adults. *Am J Clin Nutr*. 2007;85(2):377-384. doi:10.1093/ajcn/85.2.377

10. Ogawa M, Lester R, Akima H, Gorgey AS. Quantification of intermuscular and intramuscular adipose tissue using magnetic resonance imaging after neurodegenerative disorders. *Neural Regen Res*. 2017;12(12):2100-2105. doi:10.4103/1673-5374.221170

11. Le Troter A, Fouré A, Guye M, et al. Volume measurements of individual muscles in human quadriceps femoris using atlas-based segmentation approaches. *Magn Reson Mater Phy*. 2016;29:245-257. doi:10.1007/s10334-016-0535-6

12. Inhuber S, Sollmann N, Schlaeger S, et al. Associations of thigh muscle fat infiltration with isometric strength measurements based on chemical shift encoding-based water-fat magnetic resonance imaging. *Eur Radiol Exp*. 2019;3(1):45. doi:10.1186/s41747-019-0123-4

13. Tuttle LJ, Sinacore DR, Mueller MJ. Intermuscular Adipose Tissue Is Muscle Specific and Associated with Poor Functional Performance. *J Aging Res*. 2012;2012. doi:10.1155/2012/172957

14. Zacharias A, Pizzari T, English DJ, Kapakoulakis T, Green RA. Hip abductor muscle volume in hip osteoarthritis and matched controls. *Osteoarthr Cartil*. 2016;24(10):1727-1735. doi:10.1016/J.JOCA.2016.05.002

15. Belavý DL, Miokovic T, Rittweger J, Felsenberg D. Estimation of changes in volume of individual lower-limb muscles using magnetic resonance imaging (during bed-rest). *Physiol Meas*. 2011;32(1):35-50. doi:10.1088/0967-3334/32/1/003

16. Pons C, Borotikar B, Garetier M, et al. Quantifying skeletal muscle volume and shape in humans using MRI: A systematic review of validity and reliability. Nordez A, ed. *PLoS One*. 2018;13(11):e0207847. doi:10.1371/journal.pone.0207847

17. Belzunce MA, Henckel J, Fotiadou A, Di Laura A, Hart A. Automated measurement of fat infiltration in the hip abductors from Dixon magnetic resonance imaging. *Magn Reson Imaging*. 2020;72:61-70. doi:10.1016/j.mri.2020.06.019

18. Belzunce MA, Henckel J, Fotiadou A, Di Laura A, Hart A. Automated multi-atlas segmentation of gluteus maximus from Dixon and T1-weighted magnetic resonance images. *Magn Reson Mater Physics, Biol Med*. 2020;33(5):677-688. doi:10.1007/s10334-020-00839-3

19. Sakamoto M, Hiasa Y, Otake Y, et al. Automated segmentation of hip and thigh muscles in metal artifact contaminated CT using CNN. In: Fujita H, Lin F, Kim JH, eds. *International Forum on Medical Imaging in Asia 2019*. SPIE; 2019:31. doi:10.1117/12.2521440

20. Ranzini MBM, Henckel J, Ebner M, et al. Automated postoperative muscle assessment of hip arthroplasty patients using multimodal imaging joint segmentation. *Comput Methods Programs Biomed*. 2020;183:105062. doi:10.1016/j.cmpb.2019.105062

21. Tibrewala R, Pedoia V, Lee J, et al. Automatic hip abductor muscle fat fraction estimation and association with early OA cartilage degeneration biomarkers. *J Orthop Res*. 2020. doi:10.1002/JOR.24974

22. Marcon M, Berger N, Manoliu A, et al. Normative values for volume and fat content of the hip abductor muscles and their dependence on side, age and gender in a healthy population. *Skeletal Radiol*. 2016;45(4):465-474. doi:10.1007/s00256-015-2325-z

23. Grimaldi A, Richardson C, Stanton W, Durbridge G, Donnelly W, Hides J. The association between degenerative hip joint pathology and size of the gluteus medius, gluteus minimus and piriformis muscles. *Man Ther*. 2009;14(6):605-610. doi:10.1016/j.math.2009.07.004

24. Kiyoshige Y, Watanabe E. Fatty degeneration of gluteus minimus muscle as a predictor of falls. *Arch Gerontol Geriatr*. 2015;60(1):59-61. doi:10.1016/j.archger.2014.07.013

25. Inacio M, Ryan AS, Bair WN, Prettyman M, Beamer BA, Rogers MW. Gluteal muscle composition differentiates fallers from non-fallers in community dwelling older adults. *BMC Geriatr*. 2014;14(1):37. doi:10.1186/1471-2318-14-37

26. Uemura K, Takao M, Sakai T, Nishii T, Sugano N. Volume Increases of the Gluteus Maximus, Gluteus Medius, and Thigh Muscles After Hip Arthroplasty. *J Arthroplasty*. 2016;31(4):906-912.e1. doi:10.1016/j.arth.2015.10.036

27. Bley TA, Wieben O, François CJ, Brittain JH, Reeder SB. Fat and water magnetic resonance imaging. *J Magn Reson Imaging*. 2010;31(1):4-18. doi:10.1002/jmri.21895

28. Xiang QS, An L. Water-fat imaging with direct phase encoding. *J Magn Reson Imaging*. 1997;7(6):1002-1015. doi:10.1002/jmri.1880070612

29. Hernando D, Liang ZP, Kellman P. Chemical shift-based water/fat separation: A comparison of signal models. *Magn Reson Med*. 2010;64(3):811-822. doi:10.1002/mrm.22455

30. Dixon WT. Simple proton spectroscopic imaging. *Radiology*. 1984;153(1):189-194. doi:10.1148/radiology.153.1.6089263

31. Lee W, Lim SL, Xiu J, Ho M, Rumpel H. Fat-suppressed Magnetic Resonance Imaging-How to do it Perfectly.

32. Boer P. Estimated lean body mass as an index for normalization of body fluid volumes in humans. *Am J Physiol - Ren Fluid Electrolyte Physiol*. 1984;16(4). doi:10.1152/ajprenal.1984.247.4.f632

33. Maughan RJ, Watson JS, Weir J. Muscle strength and cross-sectional area in man: a comparison of strength-trained and untrained subjects. *Br J Sports Med*. 1984;18(3):149-157. doi:10.1136/bjsm.18.3.149

34. Mersmann F, Bohm S, Schroll A, Arampatzis A. Validation of a simplified method for muscle volume assessment. *J Biomech*. 2014;47(6):1348-1352. doi:10.1016/J.JBIOMECH.2014.02.007

35. Albracht K, Arampatzis A, Baltzopoulos V. Assessment of muscle volume and physiological cross-sectional area of the human triceps surae muscle in vivo. *J Biomech*. 2008;41(10):2211-2218. doi:10.1016/j.jbiomech.2008.04.020

36. Bezrukov I, Mantlik F, Schmidt H, Schölkopf B, Pichler BJ. MR-based PET attenuation correction for PET/MR imaging. *Semin Nucl Med*. 2013;43(1):45-59. doi:10.1053/j.semnuclmed.2012.08.002

37. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*. 1999;18(8):712-721. doi:10.1109/42.796284

38. Marstal K, Berendsen F, Staring M, Klein S. SimpleElastix: A User-Friendly, Multi-lingual Library for Medical Image Registration. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE; 2016:574-582. doi:10.1109/CVPRW.2016.78

39. Klein S, Staring M, Murphy K, Viergever MA, Pluim J. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Trans Med Imaging*. 2010;29(1):196-205. doi:10.1109/TMI.2009.2035616

40. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;0(NOV):863. doi:10.3389/FPSYG.2013.00863

41. Cohen J. Statistical Power Analysis for the Behavioral Sciences. *Stat Power Anal Behav Sci*. May 2013. doi:10.4324/9780203771587

42. Sjöholm T, Ekström S, Strand R, et al. A whole-body FDG PET/MR atlas for multiparametric voxel-based analysis. *Sci Rep*. 2019;9(1):1-10. doi:10.1038/s41598-019-42613-z

43. Whyte MB, Kelly P. The normal range: It is not normal and it is not a range. *Postgrad Med J*. 2018;94(1117):613-616. doi:10.1136/postgradmedj-2018-135983

44. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*. 2019;48(1):16-31. doi:10.1093/ageing/afy169

Tables

Table 1. Characteristics of the study subjects. The age values correspond to mean (standard deviation) [min-max] values, while mean (standard deviation) values are reported for the other variables.

	Male	*Female*
Subjects	N = 31	N =20
Age [years]	32.5 (8.9)	33.4 (10.1)
	[18-59]	[20-58]
Weight [kg]	74.3 (8.2)	64.8 (8.8)
Height [m]	1.80 (0.08)	1.66 (0.07)
BMI [kg/m^2]	22.9 (2.4)	23.4 (2.9)

Table 2. Mean ± SD and Median (IQR) values for volume, normalized volume and fat fraction for GMAX, GMED, GMIN and TFL.

		Volume [cm^3]		**Normalized Volume* [cm^3/Kg]**		**Fat Fraction [%]**	
		Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
GMAX	Male	734 (153)	742 (592 - 834)	12.4 (2.3)	12.5 (10.6 – 13.8)	12.1 (4.0)	10.9 (9.2 – 13.7)
	Female	577 (85)	585 (537 - 640)	12.4 (1.9)	12.8 (11.0 – 13.8)	16.6 (6.2)	15.2 (12.1 – 19.4)
	All	672 (151)	650 (555 - 775)	12.4 (2.1)	12.6 (10.8 – 13.8)	13.9 (5.4)	12.3 (10.1 – 15.9)
GMED	Male	383 (68)	387 (334 - 430)	6.5 (0.9)	6.5 (5.8 – 7.0)	9.6 (2.0)	9.1 (8.3 – 10.3)
	Female	273 (42)	266 (237 - 300)	5.9 (0.8)	5.8 (5.3 – 6.4)	11.3 (2.7)	10.4 (9.5 – 12.7)
	All	340 (80)	332 (280 - 406)	6.2 (0.9)	6.3 (5.6 – 6.7)	10.3 (2.5)	9.8 (8.6 – 11.2)
GMIN	Male	96 (18)	97 (88 - 104)	1.6 (0.2)	1.6 (1.5 – 1.7)	9.9 (2.1)	9.5 (8.6 – 10.4)
	Female	71 (11)	71 (62 - 77)	1.5 (0.1)	1.5 (1.4 – 1.6)	12.3 (2.7)	11.4 (10.4 – 14.2)
	All	86 (20)	86 (70 - 99)	1.6 (0.2)	1.6 (1.4 – 1.7)	10.9 (2.6)	10.0 (9.0 – 12.0)
TFL	Male	53 (18)	55 (39 - 63)	0.9 (0.3)	0.9 (0.7 – 1.1)	9.2 (2.7)	9.0 (7.1 – 11.2)
	Female	35 (8)	36 (28 - 42)	0.8 (0.2)	0.8 (0.6 – 0.9)	14.1 (5.3)	13.6 (10.4 – 15.8)
	All	46 (17)	43 (34 - 59)	0.8 (0.3)	0.8 (0.6 – 1.0)	11.1 (4.6)	10.2 (7.8 – 13.5)

***** Normalized volume = volume \ lean body mass

Figures

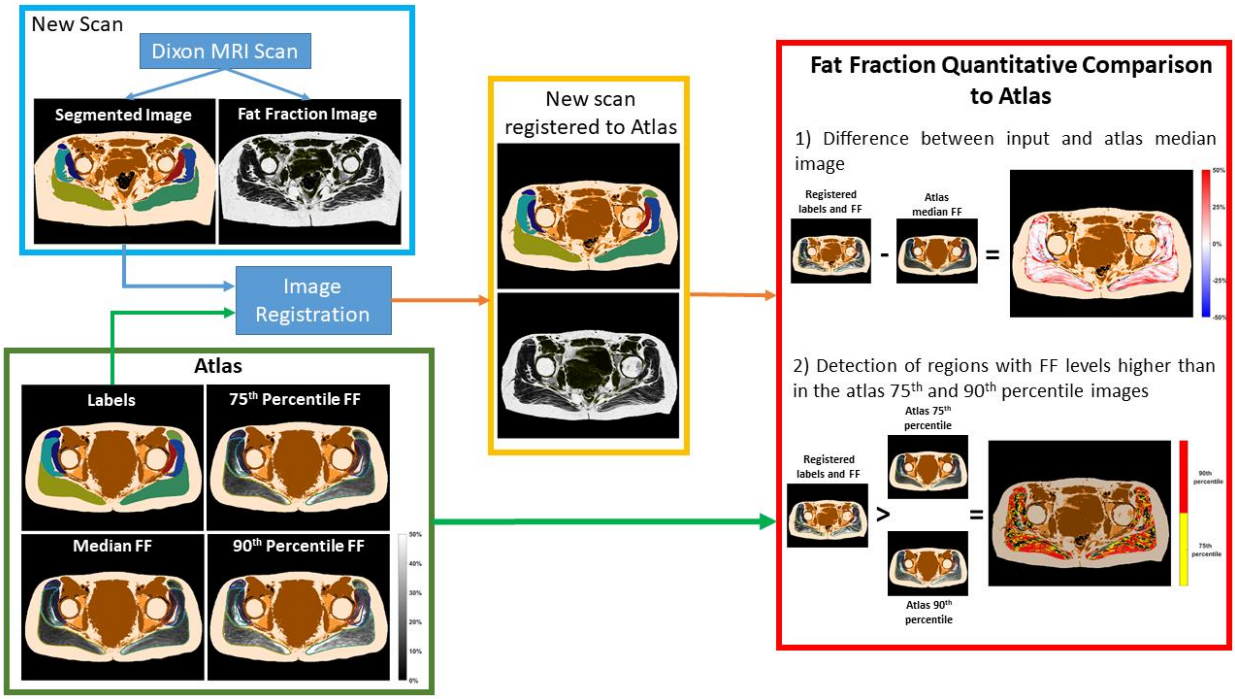

Figure 1. Processing chain for detection and localization of regions with high fat content in the hip abductors using the proposed atlas. For a new Dixon MRI scan, the in-phase image is segmented and the FF image is generated (top left). The segmented image is registered to the labels image of the atlas (bottom left) and the same transform is applied to the input FF image. The registered input (middle) is then superimposed to the FF images of the atlas (bottom left) to obtain 1) the spatial difference between the input and the median FF image of the atlas, 2) the spatial regions where the FF in the input is higher than in 75th and 90th percentile images of the atlas. The input fat fraction images are shown for its full intensity range (0-100)%, while the FF in the atlas and in the quantitative comparison are scaled to a (0-50)% range

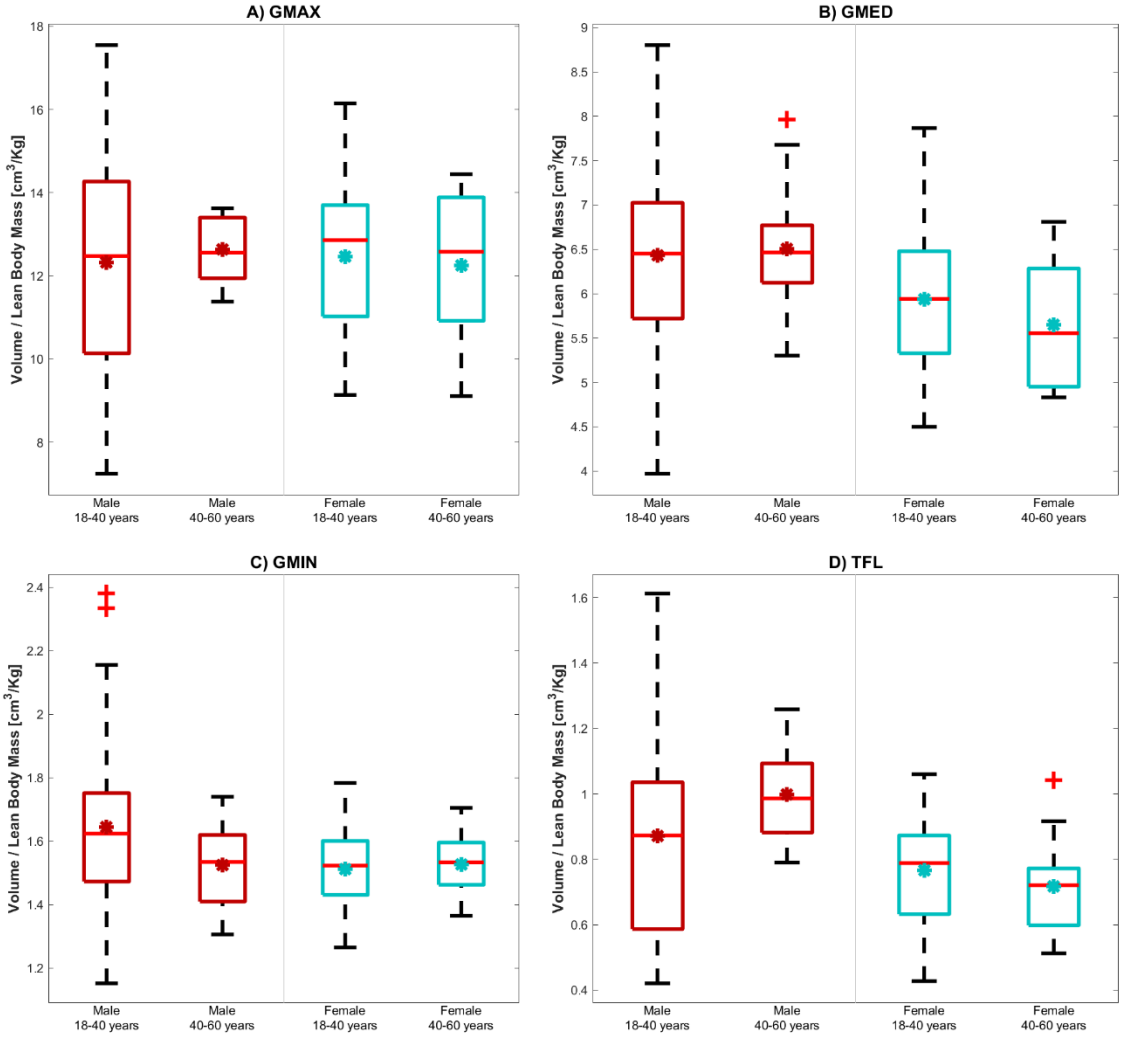

Figure 2. Boxplots of NV values of GMAX (A), GMED (B), GMIN (C) and TFL (D) for each age group and sex. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles. The mean value is also included and plotted with an *. The outliers are plotted individually.

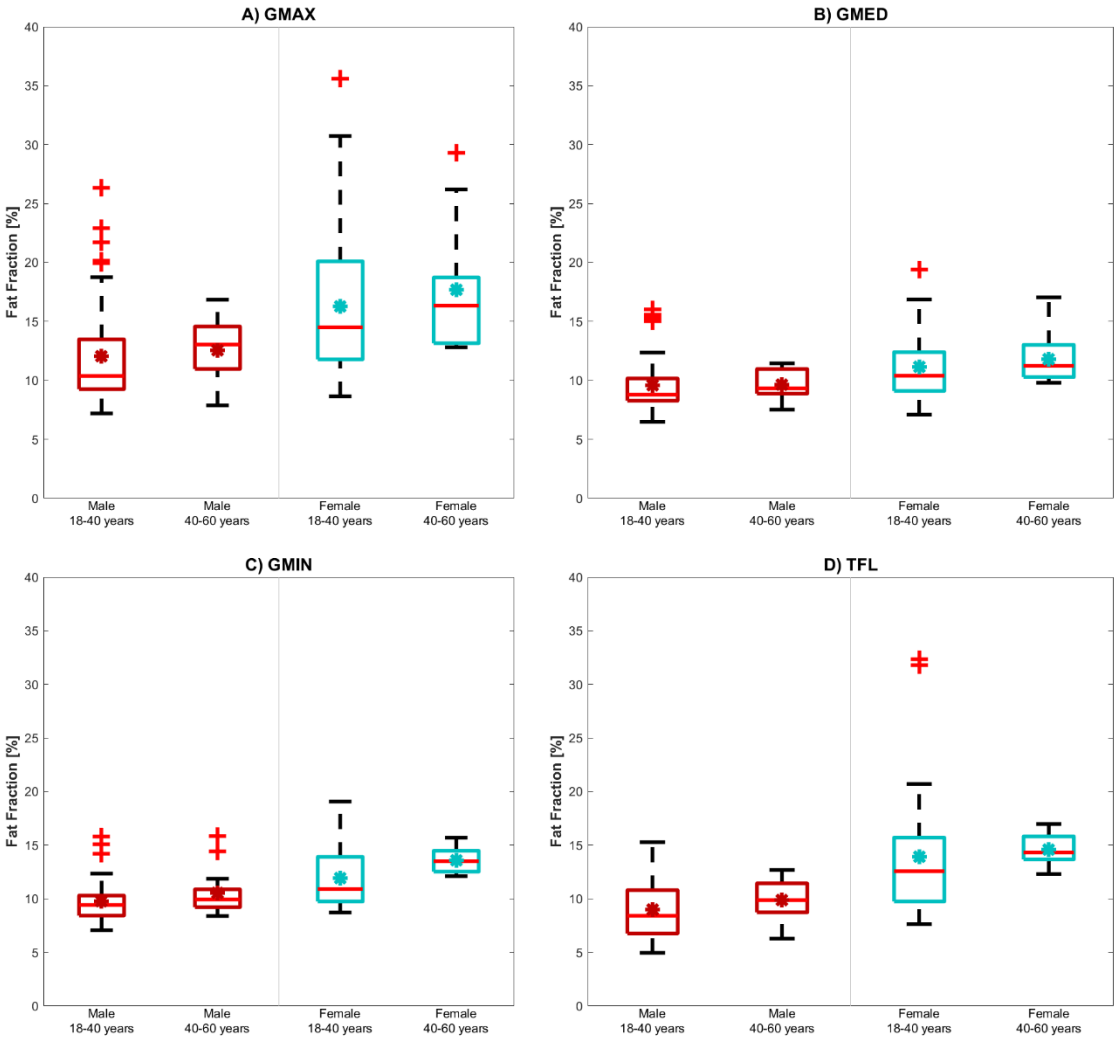

Figure 3. Boxplots of FF values of GMAX (A), GMED (B), GMIN (C) and TFL (D) for each age group and sex. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles. The mean value is also included and plotted with an *. The outliers are plotted individually.

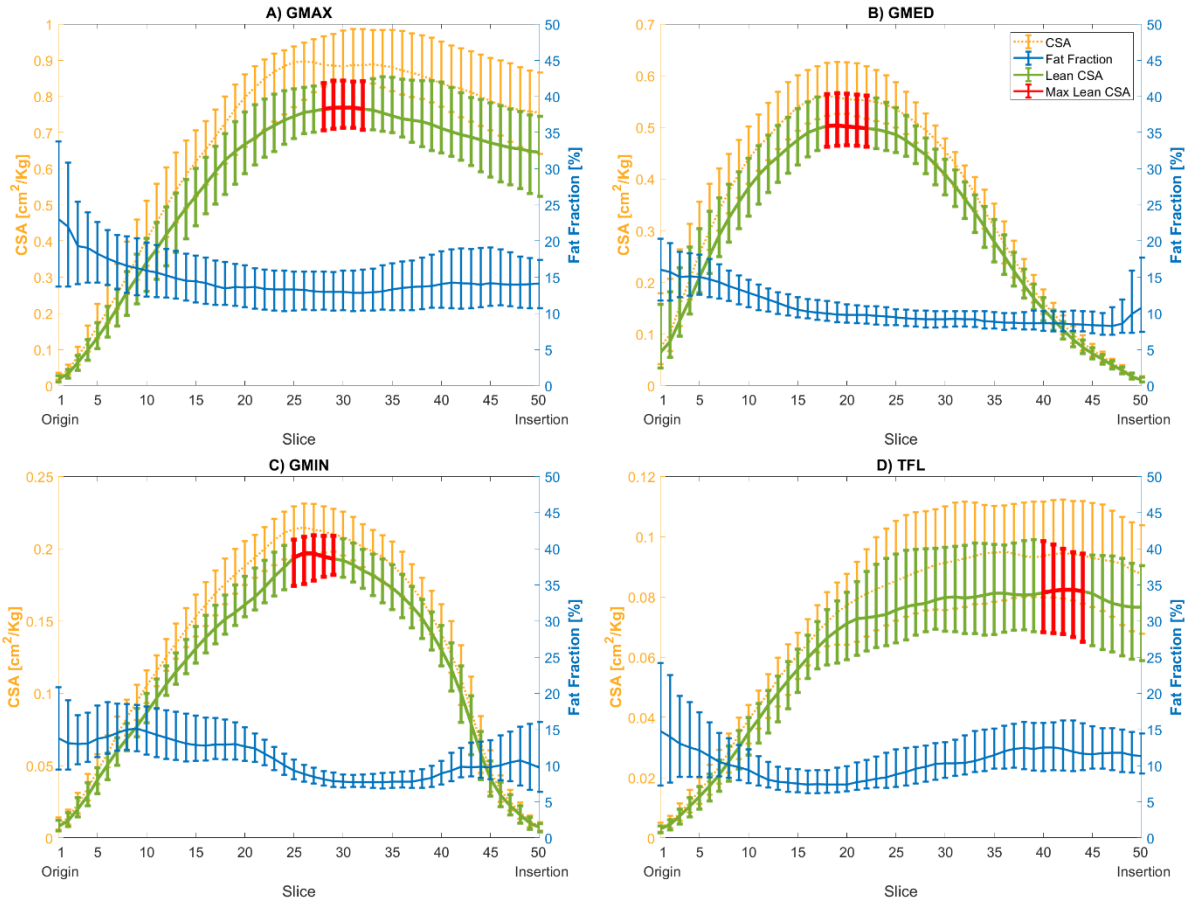

Figure 4. Axial profiles with median values and IQR error bars for normalized cross-sectional areas, fat fraction and normalized lean cross-sectional areas for A) GMAX, B) GMED, C) GMIN and D) TFL. In yellow and using the left y axis, the median normalized CSAs are plot from the origin (slice 1, most superior slice) to the insertion (slice 50, most inferior slice) of each muscle, where the error bars represent the CSA IQR in each slice for the study sample. In blue and using the right y axis, the median FF with its respective IQR error bars are plot for the same slices as for CSA. In green and using the left y axis, the profile of median and IQR error bars are plot for the normalized lean CSA, which is generated from the CSA and FF profiles. Highlighted in red, the 5 slices with highest median lean CSA.

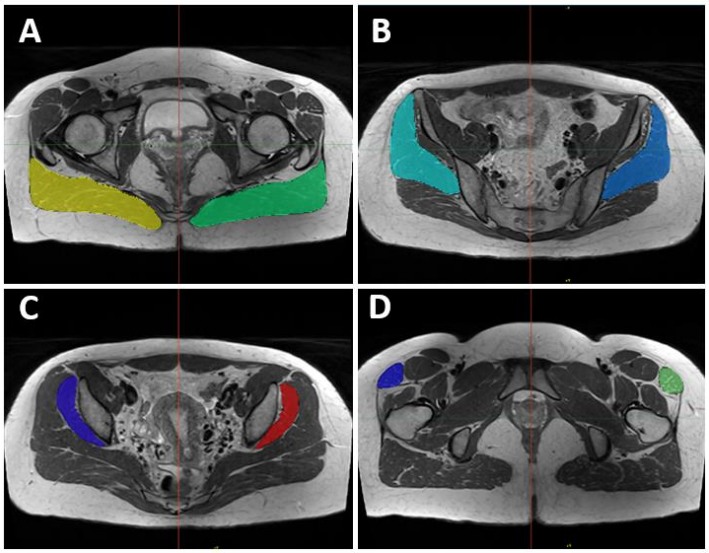

Figure 5. Example of the slices where the highest median normalized CSA values were identified for GMAX (A), GMED (B), GMIN (C) and TFL (D) in this study.

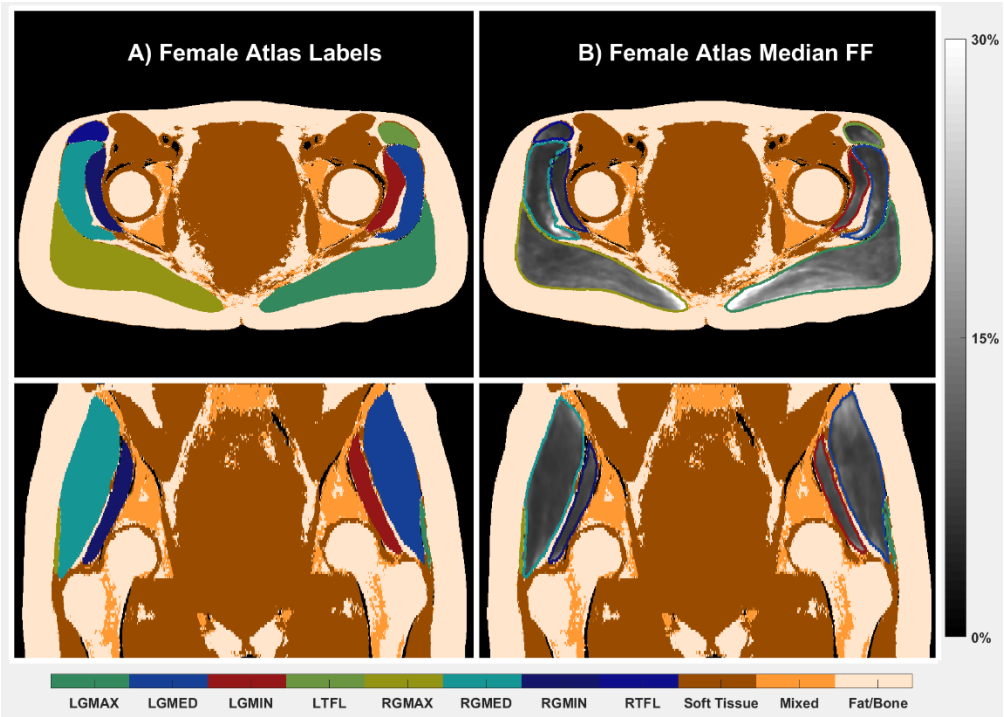

Figure 6. Female atlas. A) Axial (top) and coronal (slices) of the atlas image with labels for each of the hip abductors and tissue type from the Dixon image. The bottom colorbar identifies each label. B) Axial (top) and coronal (slices) of the FF image of the atlas that show the median FF spatial distribution for the female in our study. The colorbar on the right quantifies FF.

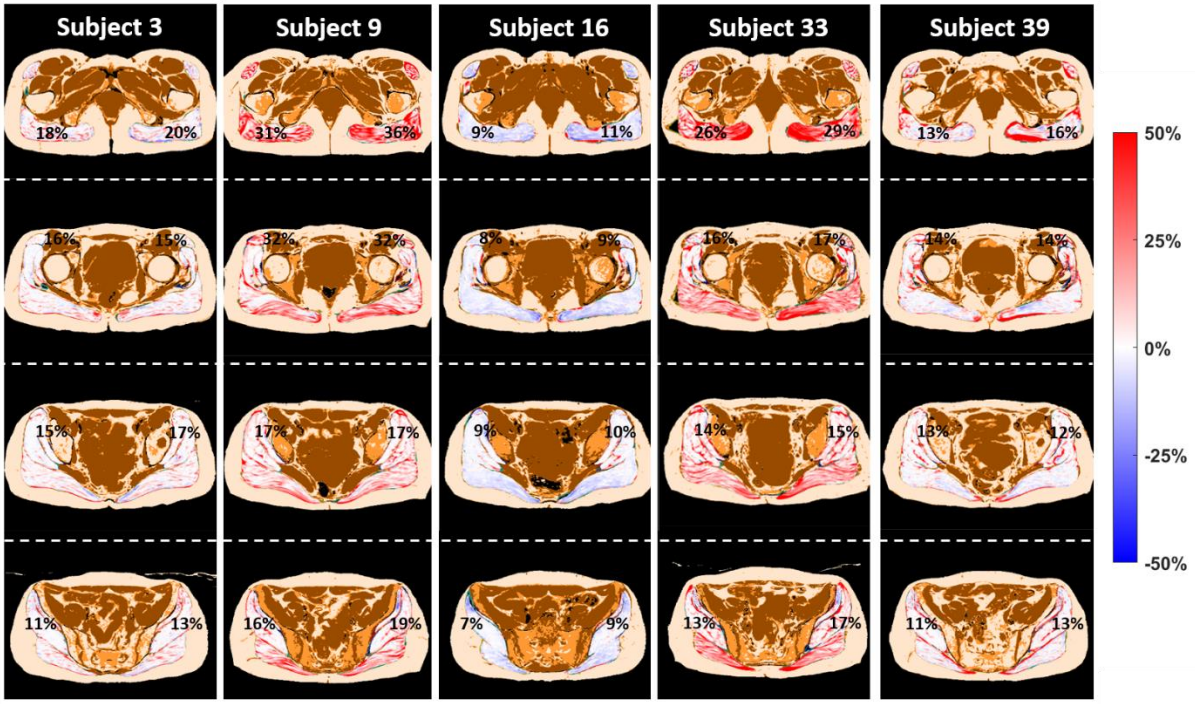

Figure 7. Analysis of 5 cases using the female atlas. Each column consists of four axial slices at different levels of the FOV, showing the FF deviation from the atlas median values at a voxel level. The colorbar shows the deviation from the median for each voxel in percentage, while in each image a legend with the mean FF value for each muscle is included. The legends correspond, from top to bottom, to GMAX, TFL, GMIN and GMED.