Precision MRI phenotyping of muscle volume and quality at a population scale

Marjola Thanaj[1], Nicolas Basty[1], Brandon Whitcher[1], Elena P. Sorokin[2], Yi Liu[2], Ramprakash Srinivasan[2], Madeleine Cule[2], E. Louise Thomas[1], and Jimmy D. Bell[1]

[1] Research Centre for Optimal Health, School of Life Sciences, University of Westminster, London, UK
[2] Calico Life Sciences LLC, South San Francisco, California, USA

Abstract

Background: Magnetic resonance imaging (MRI) enables direct measurements of muscle volume and quality, allowing for an in-depth understanding of their associations with anthropometric traits, and health conditions. However, it is unclear which muscle measurements: total muscle volume, regional measurements or measurements of muscle quality, such as intermuscular adipose tissue (IMAT) or proton density fat fraction (PDFF), are most informative to detect changes and associations with relevant health outcomes such as sarcopenia and frailty.
Methods: We developed a pipeline to automatically segment and extract image-derived phenotypes (IDPs) including total and regional muscle volumes and measures of muscle quality, and applied it to the neck-to-knee Dixon images in 44,520 UK Biobank participants. We also segmented paraspinal muscle from 2D quantitative MRI to quantify muscle PDFF and iron concentration. We performed linear regression to assess associations with anthropometric and lifestyle factors. We further applied logistic regression to investigate the association between these IDPs and sarcopenia and frailty.
Results: All muscle measurements were negatively associated with age, waist-to-hip ratio, and Townsend deprivation index while they were positively associated with body mass index, alcohol intake, hand grip strength, physical activity and were significantly higher in men. Additionally, IMAT, (corrected for muscle volume) and paraspinal muscle PDFF were significantly higher in female compared with male participants. Sarcopenia was associated with reduced muscle volume, while frailty was associated with increased IMAT and paraspinal PDFF.
Conclusions: Our fully automated method enables the quantification of muscle volumes and quality suitable for large population-based studies. While the choice of muscle measurements is important particularly when investigating associations with health conditions, most provide consistent differences relating to the detection of age, sex, body composition and lifestyle changes on muscularity.

Keywords: Muscle volume, Muscle quality, Intermuscular adipose tissue, Magnetic resonance imaging, Sarcopenia, Frailty.

Introduction

Direct population-scale measurements of muscle have been relatively limited, most studies have used measures of fat-free mass (FFM) by bioelectrical impedance analysis (BIA), or measures of lean body mass (LBM) from dual-energy X-ray absorptiometry (DXA) [1]. Unlike direct muscle measures, FFM contains non-fat components of the body, while LBM additionally contains significant and disparate contributions from organs, skin, bones, body water, and essential fat [2]. Magnetic resonance imaging (MRI) enables quantification of both muscle mass and quality, with the latter, generally related to levels of muscle fat infiltration [3,4].

Measurement of total body muscle volume and distribution by MRI has been typically limited to small cohorts, due to the cost and time-consuming requirements of image acquisition and analysis.

The lack of automated techniques led researchers to rely on the measurement of single or multiple cross-sectional areas as an index of overall muscle mass, with a single slice at the third lumbar vertebra (L3) [5], individual muscles groups such as the iliopsoas [6] or anatomical groupings such as the 'thigh muscles' [4] being among the most popular approaches. Whether proxies of muscle volume provide sufficient information to discern the overall impact of muscle volume and quality on health and ageing is yet to be fully ascertained. Similarly, there is a lack of agreement as to whether muscle volumes are best used independently or whether these should be indexed to height [7].

Ageing and morbidity-related declines in physical function, common in sarcopenia and frailty [8], are associated with a loss of muscle mass, strength and reduction in quality [9]. While sarcopenia and frailty are typically associated with ageing, it is increasingly demonstrated that multiple long-term conditions increase the risk of sarcopenia and frailty at a relatively younger age [10,11]. Recent MRI studies have shown the combination of low muscle mass and high levels of muscle fat infiltration to be a predictor of all-cause mortality, suggesting that muscle fat infiltration could contribute to the definition of sarcopenia [12]. However, much of this work has been limited to specific anatomical regions such as the thigh muscle. While the UK Biobank imaging protocol does not do full body coverage (only neck-to-knees), it is still possible to derive meaningful muscle measures including overall muscle volume within that region, specific muscle groups including iliopsoas muscle, thigh muscles, paraspinal muscles, as well as intermuscular fat content. .

Although the effect of ageing as well as sex on skeletal muscle is well known from previous studies using both DXA [7] and MRI [13], there is limited information available on how skeletal muscle tissue compartment develops across diverse populations. Previous studies using DXA reported significant differences across sex and ethnicity [14]. Here, we report both MRI muscle mass and quality measurements for 44,520 UK Biobank participants. We further assess associations between anthropometric factors and muscle volume and quality and determine which muscle IDPs are most informative and associate with relevant health conditions including sarcopenia and frailty.

Methods

Participants

A total of 44,520 UK Biobank participants were included in this analysis. Participant data from the cohort was obtained through UK Biobank Access Application number 44584. The UK Biobank has approval from the North West Multi-Centre Research Ethics Committee (REC reference: 11/NW/0382). All measurements were obtained under these ethics, adhering to relevant guidelines and regulations, with written informed consent obtained from all participants. Researchers may apply to use the UK Biobank data resources and the results generated in this study, by submitting a health-related research proposal that is in the public interest. More information may be found on the UK Biobank researchers and resource catalogue pages (https://www.ukbiobank.ac.uk).

Image processing and MRI measurements

Full details of the UK Biobank MRI abdominal protocol have previously been reported [15]. The data included in this paper focused on the neck-to-knee Dixon MRI acquisition and separate single-slice multiecho MRI acquisition for the liver. All data were analysed using our dedicated image processing pipelines for the Dixon and single-slice multiecho acquisitions, with deep learning algorithms employed to segment organs and tissue [16]. Proton density fat fraction (PDFF) and R2* were calculated from the Phase Regularized Estimation using Smoothing and Constrained Optimization (PRESCO) method [16].

Our pipeline generates more than 30 image-derived phenotypes (IDPs) from the Dixon MRI data, the subset included in this study are IDPs primarily muscle volume related, including: *total muscle* (within neck-to-knee volume), *thigh (left+right) muscle volume and iliopsoas (left+right) muscle volume*. For training the deep learning model, we utilised 108 manual annotations of total muscle and 151 annotations of the iliopsoas muscles. The Dice similarity coefficients on 20% out-of-sample testing data were 0.94 for total muscle and 0.95 for the iliopsoas muscle segmentations. The thigh muscle volumes were derived from the total muscle segmentations using anatomical landmarks from the femur. Together with these IDPs, to avoid any bias of height we also analysed a 10cm image slab, placed at the midpoint of the length of the femur on each leg. This is referred to as the *mid-thigh* region throughout the text.

In addition to the muscle IDPs several fat depots were extracted, including *abdominal subcutaneous adipose tissue (ASAT), visceral adipose tissue (VAT)* and *intermuscular adipose tissue (IMAT) volumes [16]*. We obtained a measure of intramyocellular fat stored in the paraspinal muscles, which will be referred to henceforth as PDFF to avoid confusion. We finally extracted a measure of paraspinal muscle iron concentration by transforming the R2* to iron concentration in mg/g [17]. For model training, we had 195 manual annotations of the Iterative Decomposition of Water and Fat with Echo Asymmetry and Least-Squares Estimation (IDEAL) paraspinal muscles and 187 manual annotations of the gradient echo (GRE) paraspinal muscles. The Dice similarity coefficient on 20% out-of-sample testing data was 0.81 and 0.86 for the paraspinal muscle segmentations in IDEAL and GRE acquisitions, respectively.

To account for the potential confounding effect of height on muscle and fat volumes, IDPs were indexed to height squared. Muscle volume was defined from all muscle IDPs whereas muscle quality was defined from all fat measurements including thigh IMAT volume, mid-thigh IMAT volume and paraspinal muscle PDFF however, to determine a muscle quality measurement related to IMAT, we corrected for muscle volume in the mid-thigh, represented as follows : mid-thigh IMAT/muscle ratio = (mid-thigh IMAT volume / (mid-thigh IMAT volume + mid-thigh muscle volume)) * 100.

Quality Control

IDPs with insufficient anatomical coverage were reported as missing values and a visual inspection of the MRI data was performed to determine potential extreme values in the IDPs to confirm exclusion. To ensure the full anatomical coverage of the organs a threshold was defined as follows:

- *Total muscle volume*: 5 L < total muscle < 60 L. No participants were identified above or below this threshold.
- *Thigh muscle volume*: 1 L < thigh muscle < 15 L. A total of 25 participants were identified and removed from all subsequent analyses.
- *Thigh IMAT volume*: thigh IMAT > 50 ml. One participant was identified and removed from all subsequent analyses.
- Images from all excluded participants were subsequently visually inspected to confirm the original finding.
- A total of 87 participants were excluded from the volumetric analysis of the Dixon MRI data due to missing values. We excluded 1,547 participants for the muscle and fat volume index analyses, mainly due to missing standing height data. From the single-slice multiecho data 319 participants were excluded from the PDFF and iron concentration analysis in the paraspinal muscle due to missing values.

Phenotype definitions

Anthropometric measurements including age, body mass index (BMI), waist and hip circumferences and hand grip strength (HGS) were taken at the UK Biobank imaging visit. Ethnicity was self-reported at the initial assessment visit and were categorised as follows: White (any British, Irish or other white background); Asian (any Indian, Pakistani, Bangladeshi or any other South Asian background); Black (any Caribbean, African or any other Black background); Chinese and Others (mixed ethnic background including White and Black Caribbean, White and Black African, White and Asian, and other ethnic groups). Sex was self-reported and includes those recorded by the NHS and those obtained at the initial assessment visit. The Townsend deprivation index was calculated immediately prior to participants joining the UK Biobank.

Participant's excess metabolic equivalents of task (MET) in hours per week was computed by multiplying the time in minutes spent in each activity, including walking, moderate and vigorous activity, on a typical day by the number of reported days doing the exercise and the respective MET scores using the methods previously described [18]. Excess METs scores were at 2.3, 3 and 7 for walking, moderate and vigorous activity, respectively [18]. The total physical activity MET was computed by taking the sum of walking, moderate and vigorous MET in hours per week. In our following analysis total MET will be referred as MET. All physical activity measures, alcohol intake frequency and smoking status were self-reported at the UK Biobank imaging visit.

Relevant conditions of interest including sarcopenia, sarcopenic obesity, and frailty were defined according to published criteria:

- *Sarcopenia* was defined as low muscle strength. We used HGS below 27 kg in male participants and below 16 kg in female participants [10,19]. The HGS used was that of the dominant hand. If participants reported using both hands, the mean of the right and left hand was utilised.
- *Sarcopenic Obesity* was defined as a combination of being within the two lowest quintiles of total muscle volume indexed to height squared (<6.48 L/m² in male participants and <4.88 L/m² in female participants), and within the two highest quintiles of BMI (>29.05 kg/m² in male participants and > 28.46 kg/m² in female participants) [20].
- *Frailty* was defined using the criteria adapted by Hanlon et al.[11], to be used with self-reported UK Biobank questionnaire responses which required the presence of 3 out of 5 indicators. These indicators included weight loss, exhaustion (more than half the days or nearly every day), no or only light (once a week or less) physical activity in the last 4 weeks, slow walking speed, and low HGS (Table S1). All the frailty indicators were taken at the UK Biobank imaging visit. Additional classifications included pre-frailty (1 or 2 of the 5 indicators), and not frailty (none of the indicators).

Statistical Analysis

All summary statistics, hypothesis tests, regression models and figures were performed in R software environment version 4.1.3. Figures were produced using the *ggplot2* package [21]. All descriptive characteristics are presented as means with standard deviations (SD) for quantitative variables and as percentages for categorical variables. Spearman's rank correlation coefficient (ρ) was used to assess monotonic trends between variables. The Wilcoxon rank-sum test was used to compare means between groups. A linear regression model was used to assess the associations between muscle and fat IDPs and well-known anthropometric variables that are likely to impact body composition and strength [22] including age, sex, ethnicity, BMI, waist-to-hip ratio, alcohol intake frequency, smoking status, Townsend deprivation index, dominant HGS and MET.

To test the associations between muscle volume and quality with disease status, such as sarcopenia and frailty, we constructed a linear regression model adjusted for standardised age, sex, and

binary or categorical measures of the disease status as well as standardised fat distribution adiposity including VAT and ASAT volumes separately in each model. VAT and ASAT have been previously utilised in statistical modelling for defining sarcopenia [7].

We further explored the association between muscle and fat IDPs and sarcopenia by performing a logistic regression analysis model adjusting for age, sex, ethnicity, waist-to-hip ratio, alcohol intake frequency, smoking status, Townsend deprivation index, MET and the muscle and fat IDPs including total muscle volume, thigh muscle volume, mid-thigh muscle volume, iliopsoas muscle volume, thigh IMAT volume, mid-thigh IMAT volume, mid-thigh IMAT/muscle volume, paraspinal muscle PDFF and paraspinal muscle iron concentration. We applied an ordinal logistic regression model to investigate the association between muscle and fat IDPs on frailty given it has multiple ordered categories (not frail, pre-frail, and frail). We adjusted for all the variables used in the above logistic regression model. The ordinal logistic regression model was performed using the R *MASS* package [23]. BMI and HGS were excluded from logistic models since they were used in the definition of frailty and sarcopenia, respectively. In all models iliopsoas muscle, thigh muscle, thigh IMAT volumes, and indices are shown as the sum of left and right volumes).

All anthropometric variables and IDPs were standardised. Summaries of the linear regression model are reported as either beta coefficients with 95% confidence intervals (CIs) or beta coefficients with standard errors. Summaries of the logistic regression model are reported as odds ratios (OR) with 95% CIs. The Bonferroni-corrected threshold for statistical significance was $0.05/873 = 5.7 \times 10^{-5}$.

Results

Typical images showing muscle and fat IDPs generated from each participant are shown in Figure 1. Image datasets were available from 44,520 individuals, with all required IDPs being successfully derived for each participant. Summary statistics for the study population are provided in Table 1, 51.7% were female with similar average age with male (female 63.5 ± 7.6 years and male 64.9 ± 7.8 years). The average BMI for female and male were 26.0 ± 4.7 kg/m^2 (range 13.4 to 62 kg/m^2) and 27.0 ± 3.9 kg/m^2 (range 16.4 to 58.1 kg/m^2) respectively. HGS in the dominant hand was 38.8 ± 8.9 kg in male and 24 ± 6.1 kg in female and the MET was reported as 33.6 ± 38.9 hours/week from female participants and 35.9 ± 41.9 hours/week from male participants (Table 1). The overall ethnic distribution for the cohort was 96.7% White, 1.1% Asian, 0.7% Black, 0.3% Chinese and 1% Others (see methods for details).

Figure 1 Segmentations from the UK Biobank abdominal MRI of a participant showing: **A)** the total muscle (dark red) overlaid on the coronal view of a neck-to-knee scan, **B)** the paraspinal muscle (green) in 2D quantitative MRI, **C)** the thigh muscles (yellow), thigh IMAT (blue), thigh SAT (purple) overlaid on the coronal view, and **D)** 3D renderings of the muscle segmentations for the following structures: mid-thigh muscle (yellow), mid-thigh IMAT (blue), mid-thigh subcutaneous adipose tissue (purple), paraspinal muscles (green), iliopsoas muscles (red) and total muscle (white).

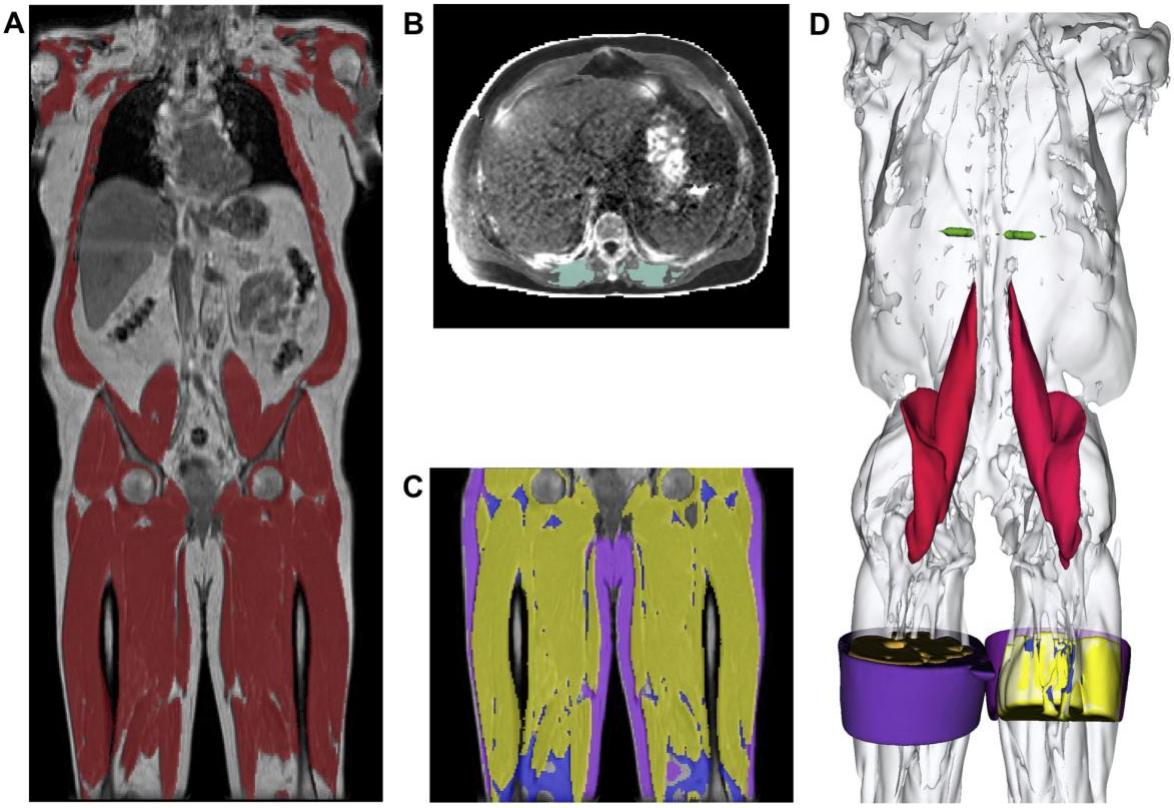


Table 1 Demographics of the participants (N=44,520), separated by sex. Values are reported as mean and standard deviation for continuous variables and counts (N) for categorical variables. Significance refers to the p-value for a Wilcoxon rank-sums test, where the null hypothesis is the medians between the two groups (male and female participants) being equal.

	Full cohort	**Female**	**Male**	**p-values**
N	44,520	23,013	21,507	-
White (N)	43,059	22,277	20,782	-
Asian (N)	476	175	301	-
Black (N)	297	164	133	-
Chinese (N)	133	80	53	-
Mixed (N)	202	132	70	-
Others (N)	438	265	173	-
Age (yrs.)	64.21 ± 7.73	63.54 ± 7.58	64.93 ± 7.83	$p < 10^{-16}$**
Weight (kg)	75.99 ± 15.10	68.88 ± 13.03	83.60 ± 13.35	$p < 10^{-16}$**
Height (cm)	169.17 ± 9.27	162.72 ± 6.25	176.06 ± 6.64	$p < 10^{-16}$**
BMI (kg/m²)	26.47 ± 4.37	26.02 ± 4.72	26.95 ± 3.90	$p < 10^{-16}$**
Waist Circumference (cm)	88.37 ± 12.69	82.78 ± 11.83	94.34 ± 10.71	$p < 10^{-16}$**

Hip Circumference (cm)	100.76 ± 8.71	100.81 ± 9.83	100.70 ± 7.34	1.4×10^{-9}**
Waist-to-Hip Ratio	0.88 ± 0.09	0.82 ± 0.07	0.94 ± 0.06	$p < 10^{-16}$**
Dominant HGS (kg)	31.17 ± 10.60	23.98 ± 6.12	38.82 ± 8.85	$p < 10^{-16}$**
Walking MET (hours/week)	12.3 ± 16.5	12.5 ± 16.9	12.1 ± 16.1	0.459
Moderate MET (hours/week)	12.7 ± 19.2	12.6 ± 18.9	12.7 ± 19.5	0.716
Vigorous MET (hours/week)	9.7 ± 17.7	8.4 ± 16.1	11.0 ± 19.3	$p < 10^{-16}$**
Total MET (hours/week)	34.70 ± 40.36	33.61 ± 38.87	35.87 ± 41.87	1.5×10^{-11}**
Townsend deprivation index	-1.88 ± 2.73	-1.83 ± 2.74	-1.93 ± 2.72	1.6×10^{-6}**
Alcohol frequency (N)				
Daily	7,585	3,118	4,467	-
1-4 times per week	24,014	11,802	12,212	-
1-3 times per month	5,130	3,081	2,049	-
Occasionally or never	7,469	4,841	2,628	-
Smoking status (N)				
Never	27,481	15,035	12,446	-
Previous	15,067	7,068	7,999	-
Current	1,518	657	861	-
VAT (L)	3.95 ± 2.31	2.82 ± 1.57	5.16 ± 2.36	$p < 10^{-16}$**
ASAT (L)	8.47 ± 4.12	9.82 ± 4.37	7.01 ± 3.27	$p < 10^{-16}$**

BMI: Body mass index; HGS: Hand grip strength; MET: Metabolic equivalents of task; VAT: Visceral adipose tissue; ASAT: Abdominal subcutaneous adipose tissue. * indicate statistically significant for $p < 0.05$, ** indicate statistically significant after Bonferroni correction ($p = 5.7 \times 10^{-5}$).

Associations with Anthropometric Factors

Total muscle in male and female participants were 21.8 ± 3.0 L and 14.0 ± 2.0 L respectively, with a mean difference between the sexes of 7.8 ± 3.6 L, or 43% (Table 2). Similar proportional sex differences were observed for all other muscle IDPs. We also observed significant differences between left and right muscle volumes for the thigh, mid-thigh, and iliopsoas muscles in both male and female participants (Table S2).

Mid-thigh IMAT levels, reflecting muscle quality, were higher in male than female participants (55.2 ± 50.3 ml vs 48.2 ± 38.4 ml, respectively) (Table 2). However, this trend was reversed when IMAT was corrected for muscle volume in the mid-thigh, represented as mid-thigh IMAT/muscle ratio, where male show a 2.0 ± 1.8 % compared to 2.5 ± 1.9 % in female participants (Table 2). We also observed a small, yet significant, difference in IMAT between left and right thigh and mid-thighs for both male and female participants ($p < 10^{-16}$, Table S2). Paraspinal muscle PDFF, was 7.8 ± 4 % in

female and 6.9 ± 3.6 % in male participants, while iron content showed a small but significant difference between sexes (Table S2).

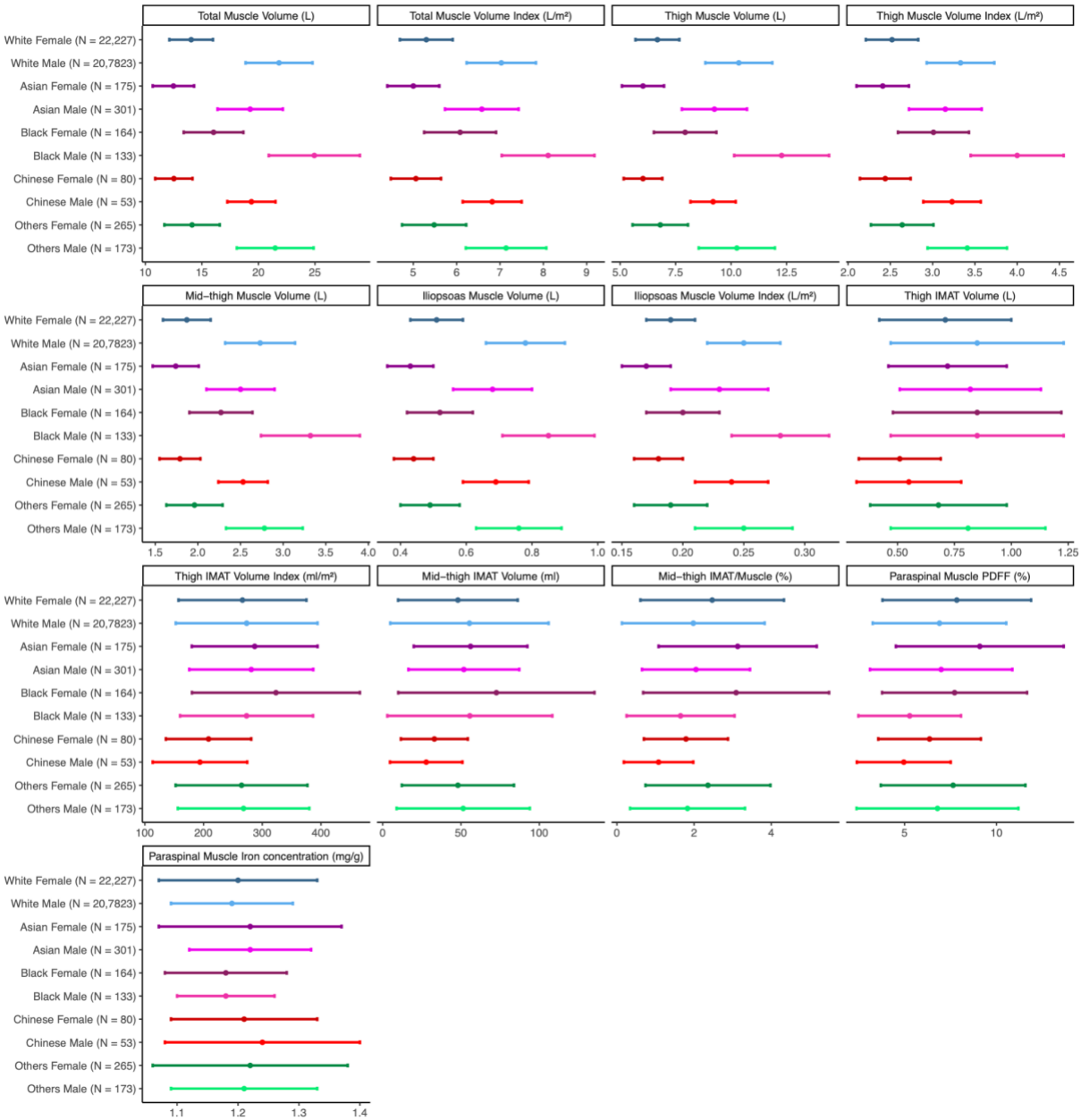
Table 2 Summary statistics for muscle volume and quality IDPs. Values are reported as mean and standard deviation. Significance refers to the p-value for a Wilcoxon rank-sums test, where the null hypothesis is the medians between the two groups (male and female participants) being equal.

	Full cohort (N = 44,520)	**Female** (N = 23,013)	**Male** (N = 21,507)	**p-values**
Muscle volume IDPs				
Total Muscle Volume (L)	17.79 ± 4.62	14.04 ± 1.95	21.80 ± 3.00	$p < 10^{-16}$**
Total Muscle Volume Index (L/m²)	6.14 ± 1.12	5.30 ± 0.62	7.03 ± 0.81	$p < 10^{-16}$**
Thigh Muscle Volume (L)	8.45 ± 2.23	6.68 ± 1.01	10.34 ± 1.53	$p < 10^{-16}$**
Thigh Muscle Volume Index (L/m²)	2.91 ± 0.54	2.52 ± 0.31	3.33 ± 0.40	$p < 10^{-16}$**
Mid-thigh Muscle Volume (L)	2.28 ± 0.55	1.87 ± 0.28	2.73 ± 0.41	$p < 10^{-16}$**
Iliopsoas Muscle Volume (L)	0.64 ± 0.17	0.51 ± 0.08	0.78 ± 0.12	$p < 10^{-16}$**
Iliopsoas Muscle Volume Index (L/m²)	0.22 ± 0.04	0.19 ± 0.02	$p < 10^{-16}$*	$p < 10^{-16}$**
Muscle quality IDPs				
Thigh IMAT Volume (l)	0.77 ± 0.34	0.71 ± 0.29	0.84 ± 0.37	$p < 10^{-16}$**
Thigh IMAT Volume Index (ml/m²)	269.82 ± 114.91	266.61 ± 109.33	273.24 ± 120.49	0.00039*
Mid-thigh IMAT Volume (ml)	51.57 ± 44.67	48.19 ± 38.40	55.19 ± 50.27	$p < 10^{-16}$**
Mid-thigh IMAT/Muscle (%)	2.24 ± 1.87	2.48 ± 1.86	1.97 ± 1.84	$p < 10^{-16}$**
Paraspinal Muscle PDFF (%)	7.38 ± 3.88	7.84 ± 4.03	6.88 ± 3.64	$p < 10^{-16}$**
Paraspinal Muscle Iron concentration (mg/g)	1.20 ± 0.12	1.20 ± 0.13	1.19 ± 0.10	$p < 10^{-16}$**

IMAT: Intermuscular adipose tissue; PDFF: Proton density fat fraction. * indicate statistically significant for $p < 0.05$, ** indicate statistically significant after Bonferroni correction ($p = 5.7 \times 10^{-5}$).

Demographics of the cohort, separated by self-reported ethnicity and sex, for muscle and fat IDPs are summarised in Tables S3 and S4, respectively. Overall, White male and female participants were aged 65.1 ± 7.8 and 63.6 ± 7.6 years old, respectively, Asian male participants were 61.3 ± 8.6 years old and Asian female participants were 60.9 ± 8.2 years old and Black male and female participants were 58.9 ± 7.5 and 59.6 ± 7.2 years old, respectively. Average total muscle volume was 21.8 ± 2.9 L in White male and 14.1 ± 1.9 L for White female participants, 19.3 ± 3 L and 12.5 ± 1.8 L for Asian male and female participants, and 24.9 ± 4.0 L and 16.0 ± 2.6 L for Black male and female participants. Mid-thigh IMAT volume was 55.4 ± 50.6 in White male and 48.0 ± 38.2 in White male participants, 51.8 ± 35.4 in Asian male and 56.1 ± 36.3 in Asian female participants, and 55.6 ± 52.7 in Black male and 72.6 ± 62.7 in Black female participants (Figure 2, Table S4).

Figure 2 Summary statistics for muscle volume and quality IDPs by sex and self-reported ethnicity. Error bars show the mean and the standard deviation.

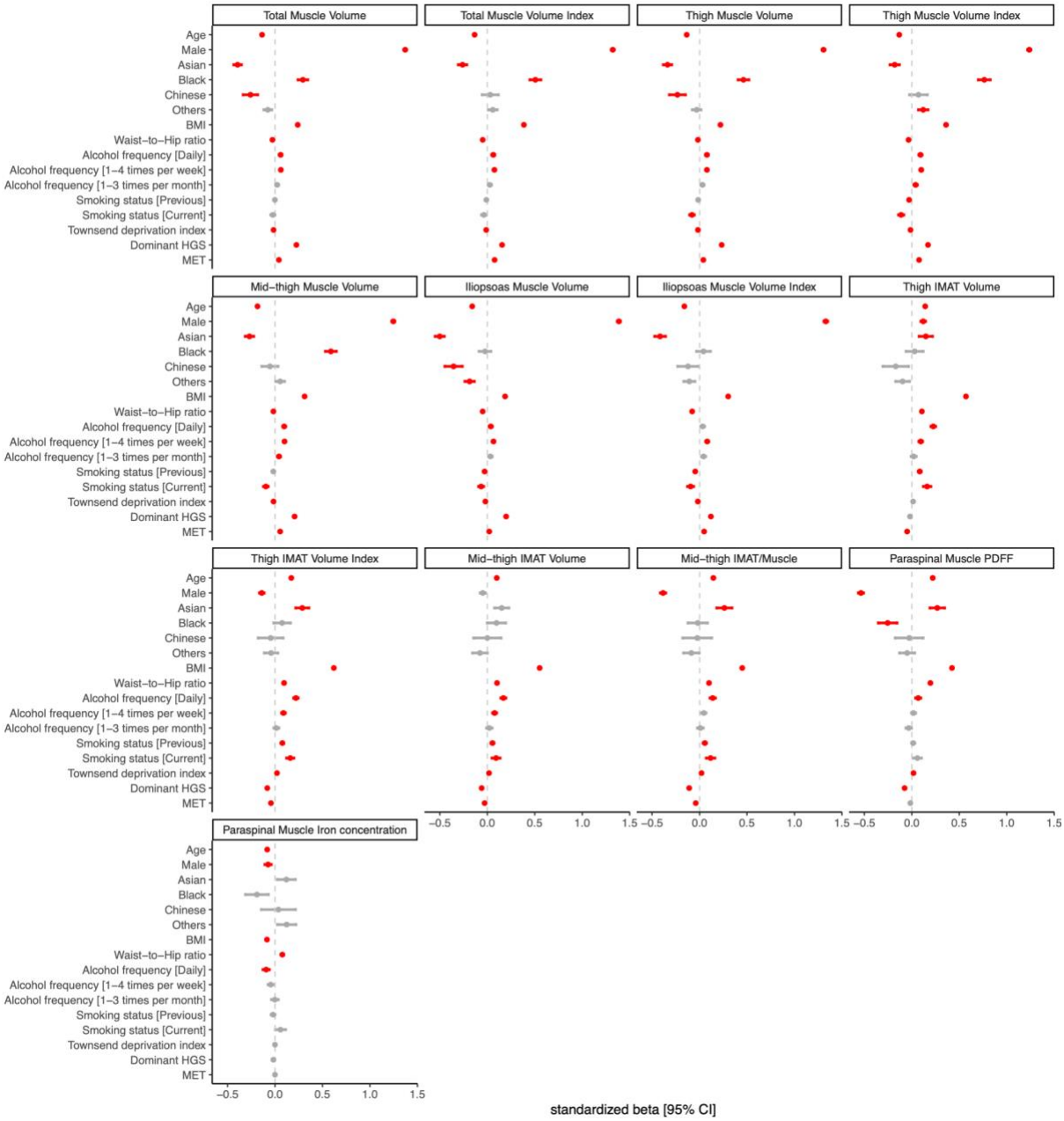


IMAT: Intermuscular adipose tissue; PDFF: Proton density fat fraction.

A multiple linear regression analysis was used to investigate the associations between muscle and fat IDPs and anthropometric variables including age, sex, ethnicity, BMI, waist-to-hip ratio, alcohol intake frequency, smoking status, Townsend deprivation index, dominant HGS and MET. Associations from this analysis are summarised in Figure 3 and Table S5. Muscle volumes were negatively associated with age, waist-to-hip ratio, and Townsend deprivation index, while thigh, mid-thigh and iliopsoas muscle IDPs were negatively associated with current smoking status (Figure 3, Table S5). Total muscle, thigh and iliopsoas muscle IDPs were positively associated with BMI, daily alcohol intake, dominant HGS and MET in hours per week and were significantly higher in men.

IDPs relating to poorer muscle quality were positively associated with increased age, BMI, waist-to-hip ratio, daily alcohol intake, smoking, and Townsend deprivation index, and negatively associated with HGS and MET. Thigh IMAT volume was significantly higher in male while, thigh IMAT volume index, mid-thigh IMAT/Muscle and paraspinal muscle PDFF were significantly lower in male (Figure 3, Table S5). Paraspinal muscle iron concentration was negatively associated with age, BMI and daily alcohol intake, but was positively associated with waist-to-hip ratio and was significantly higher in female (Figure 3, Table S5).

Figure 3 Summary of linear regression coefficients of the muscle and fat IDPs, adjusted for baseline characteristics. Standardised betas are displayed with 95% confidence intervals. Significant associations reaching the Bonferroni-corrected threshold are shown in red and non-significant associations in grey.



IMAT: Intermuscular adipose tissue; PDFF: Proton density fat fraction.

Associations between sarcopenia and frailty and muscle health

To explore the relationship between baseline characteristics and muscle IDPs in conditions typically associated with lower muscle mass and function, we identified participants who were frail, had sarcopenia, or sarcopenic obesity. From the overall cohort of 44,520 participants 3,261 (7.6%) were classified as having sarcopenia, 563 (1.3%) sarcopenic obesity and 455 (1.1%) frailty (Table 3). Representative examples of thigh MRI images from these subgroups are shown in Figure 4.

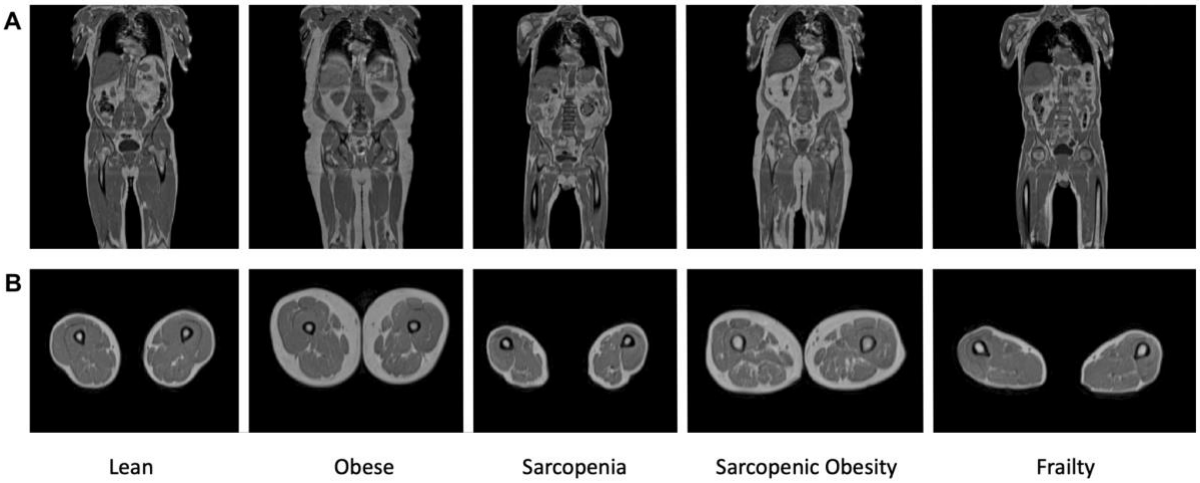
Table 3 Summary of the participants defined with sarcopenia, sarcopenic obesity and frailty.

	Full cohort (N = 44,520)	**Female** (N = 23,013)	**Male** (N = 21,507)

Sarcopenia	3,261 (7.6%)	1,668 (7.6%)	1,593 (7.7%)
Sarcopenic Obesity	563 (1.3%)	295 (0.7%)	268 (0.6%)
Frail	455 (1.1%)	273 (1.3%)	182 (0.90%)
Pre-Frail	15,686 (37.8%)	7,983 (37.5%)	7,703 (38.2%)
Not-Frail	25,349 (61.1%)	12,056 (61.3%)	12,293 (60.9%)

Note that there are overall 1,658 (914F/744M) missing data for sarcopenia, 288 (158F/130M) missing data for sarcopenic obesity, and 3,030 (1,701F/1,329M) missing data for frailty.

Figure 4 Representative samples of **A)** coronal and **B)** axial MRI views obtained from lean, obese, sarcopenic, sarcopenic obese and frail participants from the Dixon acquisition.



There was a significant correlation between dominant HGS and total muscle volume in the absence of sarcopenia in both sexes (male: ρ = 0.35; female ρ = 0.38, both $p < 5.7 \times 10^{-5}$), however this relationship was no longer significant in the presence of sarcopenia (Figure S1). We further assessed the association between sarcopenia and frailty on muscle volume using linear regression adjusted for age, sex, VAT, and ASAT (Tables S6-S18). Since the definition of sarcopenic obesity included the total muscle volume index, therefore sarcopenic obesity was excluded from further analysis. Sarcopenia and frailty were significantly associated with a reduction in all measured muscle volumes (Tables S6-S12). We also observed that reduced muscle quality was associated with sarcopenia and frailty, with increased thigh IMAT volume index, mid-thigh IMAT, mid-thigh IMAT/Muscle, and paraspinal muscle PDFF were associated with these conditions (Tables S13-S17). However, mid-thigh IMAT were not significantly associated with sarcopenia when ASAT was included in the model. Paraspinal muscle iron concentration was not significantly associated with any of the health conditions (Table S18).

To determine which IDPs were associated with sarcopenia and frailty, we created a logistic regression model, adjusting for age, sex, ethnicity, waist-to-hip ratio, alcohol intake frequency, smoking status, Townsend deprivation index and MET. To avoid associations between indexed IDPs and health conditions defined by body composition variables, we excluded indexed IDPs for this analysis. Performance was judged by the model with the lowest AIC (Akaike's Information Criterion). For sarcopenia, total muscle and total thigh volume were the most informative IDPs (Table S19), with muscle quality IDPs providing the least information. Conversely, for frailty, measures of muscle quality

including mid-thigh IMAT/Muscle, paraspinal muscle PDFF, and mid-thigh IMAT volume, were the most informative, with muscle volumes IDPs providing less information (Table S20).

Sarcopenia and frailty were positively associated with age and waist-to-hip ratio in all models (Table S19 and Table S20), while frailty was negatively associated with alcohol and MET (Table S20). Reduced muscle volume IDPs and increased muscle fat IDPs were associated with the presence of sarcopenia and frailty (Table S19 and Table S20); sarcopenia (OR = 0.361 95% CI 0.332-0.392) and frailty (OR = 0.820 95% CI 0.779-0.862) were negatively associated with total muscle volume. Mid-thigh IMAT/Muscle was positively associated with sarcopenia (OR = 1.116 95% CI 1.082-1.152) and frailty (OR = 1.328 95% CI 1.304-1.353). Paraspinal muscle iron concentration was not significantly associated with any of these conditions.

Discussion

In the present study we have utilised an automated deep-learning technique to accurately measure muscle volumes and quality in a large prospective cohort of middle-aged and older adults in the UK Biobank. We have confirmed and expanded on previous studies looking at the impact of age, sex, self-reported ethnicity, lifestyle factors, and health on muscle volume and quality [24] and have identified muscle IDPs strongly associated with sarcopenia and frailty.

Muscle volumes were significantly higher in male compared with female participants, even after correcting for height. We also found total, thigh, mid-thigh and iliopsoas muscle volumes to decrease with age, confirming earlier literature reports [25]. Previous studies have suggested that age-related changes in muscle volume begin from 27-30 years,[14,25] with a decline in muscle strength beginning at the same time, and accelerating with increasing age [26]. The cause(s) underpinning this difference is unclear as they persist even after accounting for a wide range of potential confounders, including self-reported physical activity, diet, and levels of deprivation. Further work is needed in this area to identify potential drivers associated with accelerated loss of muscle mass and quality in different age groups.

IMAT has been reported to be modulated by age, ethnicity and lifestyle factors,[3,27] findings partially confirmed in the current study. The metabolic relevance of IMAT is currently not fully understood, however while preclinical studies suggest no deleterious effects, studies in humans have linked it to adverse metabolic health [27]. IMAT increases in response to obesity and sedentary behaviours, with elevated levels linked to a decline in insulin sensitivity [28]. It has also been suggested that since IMAT secretes proteins that modify muscle extracellular matrix, it may promote muscle fibrosis and consequently the development/progression of sarcopenia [28]. Given the potential role that muscle quality appears to play in multiple conditions, including NAFLD [29], neuromuscular diseases [30], whole-muscle extensibility [31], and poor functional recovery following whiplash [32], its accurate measurement and presentation has become essential for our understanding of muscle and fat in the development and progression of many clinical conditions.

Here we show total thigh IMAT content was significantly higher in male compared to female participants, however when presented as the mid-thigh IMAT/muscle ratio, this relationship was reversed, with higher levels of mid-thigh IMAT/muscle found in female participants. Similarly, paraspinal muscle PDFF was significantly higher in female participants. This clearly highlights the impact of the type of measurement on possible outcomes. Previously studies have reported higher whole-body and thigh IMAT levels in men [3,27], but these were generally not corrected for muscle volume. Although, correcting for muscle mass negated reported sex differences in IMAT [33].

In contrast to previous studies [34], we found that paraspinal muscle PDFF increased with both age and BMI, which may be a reflection of the larger cohort size and greater age range than has previously been employed for these measurements. Paraspinal muscle PDFF is used as a prognostic

biomarker of conditions including neuromuscular disorders, intervertebral disc disease, back pain, osteoporosis, and sarcopenia [34]. We also found, in agreement with previous reports, higher levels of paraspinal muscle PDFF in female compared with male participants [34], which may correspond with higher reported prevalence and severity of lumbar disease in females [35].

Paraspinal muscle iron was higher in female compared with male participants, although the magnitude of this difference was small. Previous reports have suggested that muscle iron levels are higher in men [36], but differences in methodology (MRI vs biopsy), and muscle type (paraspinal vs vastus lateralis), make a direct comparison difficult. Some studies have suggested that muscle iron is associated with body fat mass and is increased in obesity [37], however, we found a negative association with BMI and age.

The prevalence of sarcopenia (7.6%) and frailty (1.1%) were relatively low in our cohort, in agreement with reports that the UK Biobank cohort is a comparatively healthy population [38]. We sought to determine which of our measured muscle IDPs were associated with these conditions, and which would potentially provide the best prognostic tool for identifying this population. We defined sarcopenia solely based on low muscle strength [10], not including reduced muscle mass/quality and/or reduced physical performance, and found this definition of sarcopenia was associated with reduced muscle mass and increased levels of mid-thigh IMAT/muscle and paraspinal PDFF. While for sarcopenia, the muscle volume measurements with the greatest body coverage such as total muscle or the entire thigh were the most informative measurements, for frailty markers of muscle quality, particularly paraspinal muscle PDFF, mid-thigh IMAT/Muscle, and mid-thigh IMAT were the most informative IDPs. This suggests that different IDPs may have different diagnostic values for different conditions.

Our study is not without limitations. The UK Biobank is a large cross-sectional study that is subject to selection bias with a “healthier” than the wider UK population, excludes younger participants and potentially more severe clinical cases [39]. Furthermore, as the UK Biobank study includes predominantly White population, future work is needed to further investigate imbalance population data. Also, given the UK Biobank neck-to-knee acquisition, measurement of whole-body muscle volume could not be achieved, thus limiting direct comparisons with some literature values [40]. We estimate that the current protocol lacks c.a. 9.75% of total muscle, corresponding to both arms and lower calves. However, as our findings were robust to the selection of muscle volume or quality measure, we consider it unlikely that results with a whole body measure would be dramatically different.

Conclusions

We have presented a fully automated method that enables the measurement of muscle volumes and quality suitable for application to large population-based studies. We have reported that while the choice of muscle measurement may have an impact on the ability to identify disease, most provide consistent differences relating to the detection of the impact of age, sex, and clinical conditions on muscularity.

Acknowledgments
This research has been conducted using the UK Biobank Resource under Application Number 44584.

Author Contributions
J.D.B., E.L.T., M.T. and M.C. conceived the study. J.D.B., B.W., E.L.T., N.B. and M.T. designed the study. M.T., N.B., B.W., Y.L., E.P.S. and M.C. implemented the methods and performed the data analysis. M.T. defined the disease categories. M.T. performed the statistical analysis. E.L.T., B.W., M.T., J.D.B., R.S. and N.B. drafted the manuscript. All authors read and approved the manuscript.

Funding
This work was funded by Calico Life Sciences LLC.

Conflict of Interests
M.C., R.S, and E.P.S. are employees of Calico Life Sciences LLC. Y.L. is a former employee of Calico Life Sciences LLC. M.T., N.B., B.W., J.D.B. and E.L.T. declare no competing interests.

Ethical Approval
The UK Biobank has approval from the North West Multi-Centre Research Ethics Committee (REC reference: 11/NW/0382), and obtained written informed consent from all participants prior to the study. All methods were performed in accordance with the relevant guidelines and regulations as presented by the relevant authorities, including the Declaration of Helsinki.

References

1. Zillikens MC, Demissie S, Hsu Y-H, Yerges-Armstrong LM, Chou W-C, Stolk L, et al. Large meta-analysis of genome-wide association studies identifies five loci for lean body mass. Nat Commun. 2017;8:80.

2. Scafoglieri A, Clarys JP. Dual energy X-ray absorptiometry: gold standard for muscle mass? J. Cachexia Sarcopenia Muscle. 2018. p. 786–7.

3. Gallagher D, Kuznia P, Heshka S, Albu J, Heymsfield SB, Goodpaster B, et al. Adipose tissue in muscle: a novel depot similar in size to visceral adipose tissue. Am J Clin Nutr. 2005;81:903–10.

4. Linge J, Borga M, West J, Tuthill T, Miller MR, Dumitriu A, et al. Body Composition Profiling in the UK Biobank Imaging Study. Obesity . 2018;26:1785–95.

5. Schweitzer L, Geisler C, Pourhassan M, Braun W, Glüer C-C, Bosy-Westphal A, et al. What is the best reference site for a single MRI slice to assess whole-body skeletal muscle and adipose tissue volumes in healthy adults? Am J Clin Nutr. 2015;102:58–65.

6. Fitzpatrick JA, Basty N, Cule M, Liu Y, Bell JD, Thomas EL, et al. Large-scale analysis of iliopsoas muscle volumes in the UK Biobank. Sci Rep. 2020;10:20215.

7. Linge J, Heymsfield SB, Dahlqvist Leinhard O. On the Definition of Sarcopenia in the Presence of Aging and Obesity-Initial Results from UK Biobank. J Gerontol A Biol Sci Med Sci. 2020;75:1309–16.

8. Ungar A, Marchionni N. Cardiac Management in the Frail Elderly Patient and the Oldest Old. Springer; 2017.

9. Sizoo D, de Heide LJM, Emous M, van Zutphen T, Navis G, van Beek AP. Measuring Muscle Mass and Strength in Obesity: a Review of Various Methods. Obes Surg. 2021;31:384–93.

10. Dodds RM, Granic A, Robinson SM, Sayer AA. Sarcopenia, long-term conditions, and multimorbidity: findings from UK Biobank participants. J Cachexia Sarcopenia Muscle. 2020;11:62–8.

11. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Health. 2018;3:e323–32.

12. Linge J, Ekstedt M, Dahlqvist Leinhard O. Adverse muscle composition is linked to poor

functional performance and metabolic comorbidities in NAFLD. JHEP Rep. 2021;3:100197.

13. Yang YX, Chong MS, Lim WS, Tay L, Yew S, Yeo A, et al. Validity of estimating muscle and fat volume from a single MRI section in older adults with sarcopenia and sarcopenic obesity. Clin Radiol. 2017;72:427.e9–427.e14.

14. Silva AM, Shen W, Heo M, Gallagher D, Wang Z, Sardinha LB, et al. Ethnicity-related skeletal muscle differences across the lifespan. Am J Hum Biol. 2010;22:76–82.

15. Littlejohns TJ, Holliday J, Gibson LM, Garratt S, Oesingmann N, Alfaro-Almagro F, et al. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. Nat Commun. 2020;11:2624.

16. Liu Y, Basty N, Whitcher B, Bell JD, Sorokin EP, van Bruggen N, et al. Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. Elife [Internet]. 2021;10. Available from: http://dx.doi.org/10.7554/eLife.65554

17. Wood JC, Enriquez C, Ghugre N, Tyzka JM, Carson S, Nelson MD, et al. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. Blood. 2005;106:1460–5.

18. O'Donnell J, Smith-Byrne K, Velardo C, Conrad N, Salimi-Khorshidi G, Doherty A, et al. Self-reported and objectively measured physical activity in people with and without chronic heart failure: UK Biobank analysis. Open Heart. 2020;7:e001099.

19. Dodds RM, Syddall HE, Cooper R, Benzeval M, Deary IJ, Dennison EM, et al. Grip strength across the life course: normative data from twelve British studies. PLoS One. 2014;9:e113637.

20. Stenholm S, Harris TB, Rantanen T, Visser M, Kritchevsky SB, Ferrucci L. Sarcopenic obesity: definition, cause and consequences. Curr Opin Clin Nutr Metab Care. 2008;11:693–700.

21. Wilkinson L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H [Internet]. Biometrics. 2011. p. 678–9. Available from: http://dx.doi.org/10.1111/j.1541-0420.2011.01616.x

22. Farmer RE, Mathur R, Floriaan Schmidt A, Bhaskaran K, Fatemifar G, Eastwood SV, et al. Associations Between Measures of Sarcopenic Obesity and Risk of Cardiovascular Disease and Mortality: A Cohort Study and Mendelian Randomization Analysis Using the UK Biobank [Internet]. Journal of the American Heart Association. 2019. Available from: http://dx.doi.org/10.1161/jaha.118.011638

23. Venables WN, Ripley BD. Modern Applied Statistics with S-PLUS. Springer Science & Business Media; 2013.

24. Kalyani RR, Corriere M, Ferrucci L. Age-related and disease-related muscle loss: the effect of diabetes, obesity, and other diseases. Lancet Diabetes Endocrinol. 2014;2:819–29.

25. Gallagher D, Visser M, De Meersman RE, Sepúlveda D, Baumgartner RN, Pierson RN, et al. Appendicular skeletal muscle mass: effects of age, gender, and ethnicity. J Appl Physiol. 1997;83:229–39.

26. Clement FJ. Longitudinal and cross-sectional assessments of age changes in physical strength as related to sex, social class, and mental ability. J Gerontol. 1974;29:423–9.

27. Sparks LM, Goodpaster BH, Bergman BC. The Metabolic Significance of Intermuscular Adipose Tissue: Is IMAT a Friend or a Foe to Metabolic Health? Diabetes. 2021;70:2457–67.

28. Sachs S, Zarini S, Kahn DE, Harrison KA, Perreault L, Phang T, et al. Intermuscular adipose

tissue directly modulates skeletal muscle insulin sensitivity in humans. Am J Physiol Endocrinol Metab. 2019;316:E866–79.

29. Bot D, Droop A, Lucassen CJ, van Veen ME, van Vugt JLA, Shahbazi Feshtali S, et al. Both muscle quantity and quality are predictors of waiting list mortality in patients with end-stage liver disease. Clin Nutr ESPEN. 2021;42:272–9.

30. Dahlqvist JR, Widholm P, Leinhard OD, Vissing J. MRI in Neuromuscular Diseases: An Emerging Diagnostic Tool and Biomarker for Prognosis and Efficacy. Ann Neurol. 2020;88:669–81.

31. Giambini H, Hatta T, Gorny KR, Widholm P, Karlsson A, Leinhard OD, et al. Intramuscular fat infiltration evaluated by magnetic resonance imaging predicts the extensibility of the supraspinatus muscle. Muscle Nerve. 2018;57:129–35.

32. Abbott R, Peolsson A, West J, Elliott JM, Åslund U, Karlsson A, et al. The qualitative grading of muscle fat infiltration in whiplash using fat and water magnetic resonance imaging. Spine J. 2018;18:717–25.

33. Kim JE, Dunville K, Li J, Cheng JX, Conley TB, Couture CS, et al. Intermuscular Adipose Tissue Content and Intramyocellular Lipid Fatty Acid Saturation Are Associated with Glucose Homeostasis in Middle-Aged and Older Adults. Endocrinol Metab (Seoul). 2017;32:257–64.

34. Schlaeger S, Inhuber S, Rohrmeier A, Dieckmeyer M, Freitag F, Klupp E, et al. Association of paraspinal muscle water-fat MRI-based measurements with isometric strength measurements. Eur Radiol. 2019;29:599–608.

35. MacLean MA, Touchette CJ, Han JH, Christie SD, Pickett GE. Gender differences in the surgical management of lumbar degenerative disease: a scoping review. J Neurosurg Spine. 2020;1–18.

36. Ryan BJ, Foug KL, Gioscia-Ryan RA, Ludzki AC, Ahn C, Schleh MW, et al. Skeletal muscle ferritin abundance is tightly related to plasma ferritin concentration in adults with obesity. Exp Physiol. 2020;105:1808–14.

37. Moreno-Navarrete JM, Blasco G, Xifra G, Karczewska-Kupczewska M, Stefanowicz M, Matulewicz N, et al. Obesity Is Associated With Gene Expression and Imaging Markers of Iron Accumulation in Skeletal Muscle. J Clin Endocrinol Metab. 2016;101:1282–9.

38. Lyall DM, Quinn T, Lyall LM, Ward J, Anderson JJ, Smith DJ, et al. Quantifying bias in psychological and physical health in the UK Biobank imaging sub-sample. Brain Commun. Oxford Academic; 2022;4:fcac119.

39. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47:226–35.

40. Karlsson A, Rosander J, Romu T, Tallberg J, Grönqvist A, Borga M, et al. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. J Magn Reson Imaging. 2015;41:1558–69.