Weight-related changes in MRI-derived measures of body composition and liver steatosis: a large-scale analysis for obesity trial design

	Legend	Page
Supplementary Methods		2
Supplementary Figure 1	Correlations between MRI body composition metrics at baseline. The matrix displays Spearman's correlation coefficients with color intensity indicating strength and direction. Asterisks below each value denote statistical significance after FDR correction (*p<0.05, **p<0.01, ***p<0.001).	5
Supplementary Figure 2	Correlations between relative % change in MRI body composition metrics during the follow-up period. The matrix displays Spearman's correlation coefficients with color intensity indicating strength and direction. Asterisks below each value denote statis-tical significance after FDR correction (*p<0.05, **p<0.01, ***p<0.001).	6
Supplementary Figure 3	Correlations between MRI-derived abdominal skeletal muscle (SM) volume and DXA-derived trunk lean mass (A) and total lean mass (B) at baseline across the whole cohort.	7
Supplementary Table 1	Absolute change in MRI metrics across the whole cohort and individuals whose weight remained stable, who lost weight or gained weight over the follow-up period.	8
Supplementary Table 2	Relative change in MRI metrics across individuals who were not on medications for hypertension and whose weight remained stable, who lost weight or gained weight over the follow-up period.	9
Supplementary Table 3	Relative change in MRI metrics in males whose weight remained stable, who lost weight or gained weight over the follow-up period.	10
Supplementary Table 4	Relative change in MRI metrics in females whose weight remained stable, who lost weight or gained weight over the follow-up period.	11
Supplementary Table 5	Expected changes per each 5% weight reduction in individuals with overweight or obesity based on linear regression analysis.	12
Supplementary References		13

Supplementary Methods

Data collection

Clinical and biochemical data

Prior history of diabetes, hypertension, recent cancer (within 2 years of imaging visit) and self-reported medications, were collected at the imaging visits. From hospitalisation data, primary care data and self-reports, diabetes was defined as ICD-10 E11 and E14. Hypertension defined as ICD-10 I10. From hospitalisation data, cancer was defined as ICD-10: C00-C97, D00-D09; ICD-9: 140-208, 230-234.

Body composition measurements from MRI

Data preprocessing of body composition measurements from MRI

Body composition measurements were derived using fully automated volumetric and semi-automated single-slice area analyses at the 3rd lumbar vertebra (L3) level, as described previously [1]. For each Dixon 3D SPGR water/fat volume pair, signal fat fraction (SFF) was calculated by, firstly correcting scanner-derived fat and water on a voxel-wise basis to remove bias introduced by the difference in T1 relaxation times [2]. Correction factors used T1 values of 1042 ms for water (representing muscle) and 385 ms for fat [3,4]. Secondly, the value of each voxel was computed as the fraction of the corrected fat signal relative to the sum of the corrected water and fat signals in that voxel.

Each set of water, fat and SFF volumes underwent a trilinear resampling to output a consistent voxel resolution of 2x2x3 mm, where 3mm was the resolution along the inferior-superior axis. N4 bias correction [5] was applied to fat and water volumes.

For each image type a unique whole-body space was created to extend all volumes into a single matrix. Individual volumes were sequentially added to this space from the most inferior to the most superior volume. If there was an overlap when adding a new volume, overlapping

voxel values were calculated as a weighted mean of the values from each overlapping volume, where the weighting is based on the distances between the voxel and the edge of the overlapping volumes. Signal intensity between adjacent volumes was normalized by mapping the histogram of the more inferior block to the one of the adjacent superior block. The mapping was applied to the inferior block to ensure a consistent contrast across the whole-body. This process was repeated until all volumes were contained within the whole-body volume. N4 and signal normalization was not applied to SFF volumes to avoid bias in tissue characterization.

Abdominal volumetric analysis of body composition measurements from MRI

For volumetric analysis, the region from the center of the T9 vertebra to the top of the lower of the two femoral heads was selected. To automatically identify the region limits, vertebrae and femurs were segmented in the whole-body fat volume using a nnU-Net [6]. Each independent vertebra was given a unique label from inferior to superior, starting with L5 upwards. Each femur was given a unique label, and the top of the lower femoral head was identified as the more inferior of the highest labelled slice for the two femurs. The identified slices were used to crop the fat, water and SFF volumes to the analysis region. The cropped fat and water volumes were then used as input to a 3D neural network that segmented SAT, muscle tissue, internal cavity, and bone.

The dataset included instances of fat-water swapping, which would negatively impact the fat fraction calculation, segmentation performance, and metric accuracy. Cases with fat swaps were detected using a ResNet [7] and manually corrected by re-labelling the affected blocks before the whole-body volume creation, avoiding swap artefacts.

Single-slice cross-sectional area analysis of body composition measurements from MRI

For single-slice analysis, axial slices passing through the centre of mass of the L3 vertebra were automatically extracted from the whole-body volumes for measurement. We applied a

neural network model, different from the one used in the volumetric analysis, that segmented all vertebrae within the volume as a single vertebrae class. Morphological operations were used subsequently to label individual vertebra, identifying the L3 counting vertebrae from the sacrum upwards. Within the extracted axial slice at the centre of L3 vertebra, ROI were segmented with two sequential neural networks. Firstly, a U-Net classified voxels into eight classes: SAT, abdominal cavity, and abdominal SM (including coarse segmentations of the right and left back muscles, right and left psoas muscles, and abdominal wall muscles). Secondly, an additional U-Net was applied to a focused sub-region of the image, refining psoas segmentations.

Two rounds of manual quality control (QC) were performed. The first QC step assessed the L3 centre of mass placement, rejecting cases where L3 was poorly visible or repositioning the landmark as needed. The second QC step involved a visual review of every 2D axial segmentation at the centre of L3, rejecting cases with segmentation errors or unclear tissue boundaries.

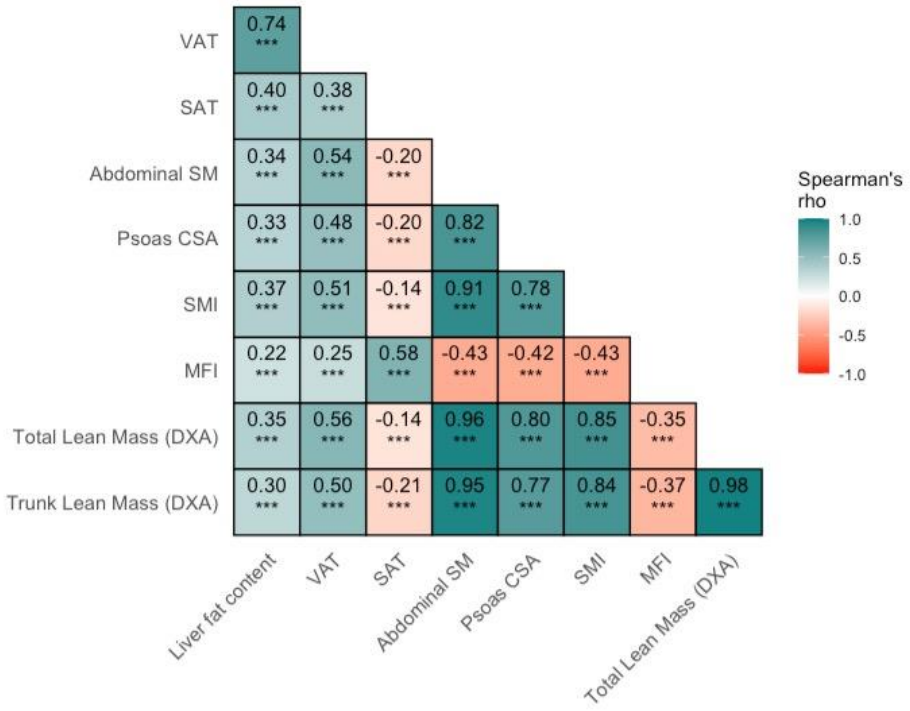
Extraction of body composition measurements from MRI

Fat and muscle tissues were quantified by filtering segmented voxels based on SFF values. Fat tissue was defined as having SFF values $^{3}0.5$, and muscle tissue as having SFF values <0.5, ensuring that any over-segmented voxels from adjacent tissues were excluded. Volume and cross sectional area (CSA) metrics were calculated based on the number of voxels and pixels, respectively. MFI was computed as the mean SFF value of the muscle tissue voxels.

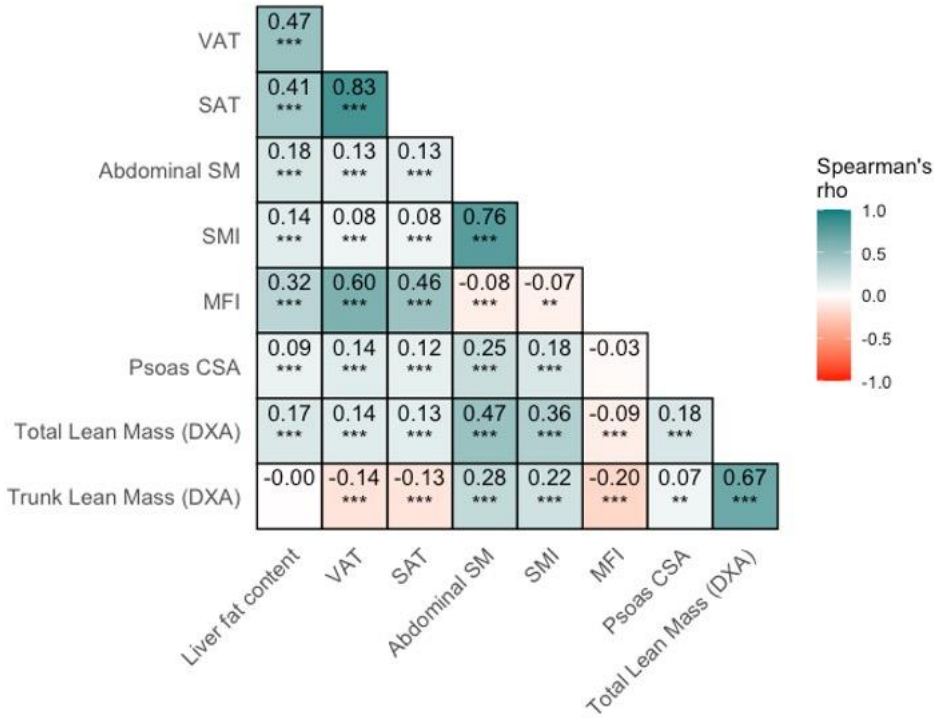
Statistical analysis

Individuals with weight changes exceeding 15% in either direction were excluded from group analyses, as this combined category (>15% weight gain or loss) comprised only 42 participants in total, providing insufficient sample size for statistical comparisons.

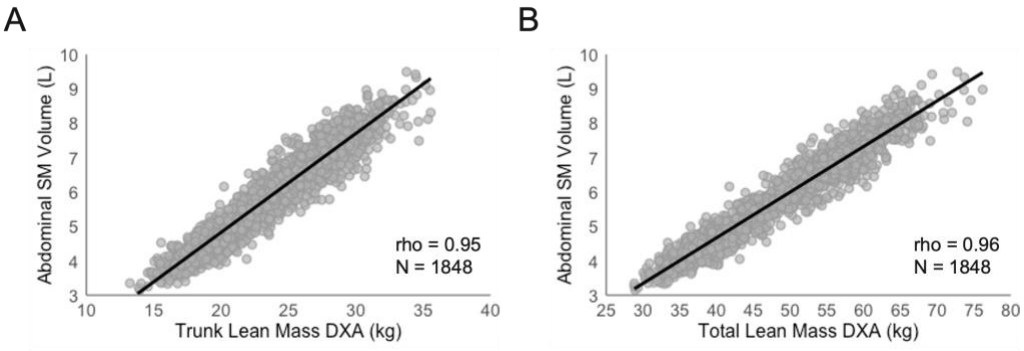
Supplementary Figure 1. Correlations between MRI body composition metrics at baseline. The matrix displays Spearman's correlation coefficients with color intensity indicating strength and direction. Asterisks below each value denote statistical significance after FDR correction (*p<0.05, **p<0.01, ***p<0.001).



Supplementary Figure 2. Correlations between relative % change in MRI body composition metrics during the follow-up period. The matrix displays Spearman's correlation coefficients with color intensity indicating strength and direction. Asterisks below each value denote statis-tical significance after FDR correction (*p<0.05, **p<0.01, ***p<0.001).



Supplementary Figure 3. Correlations between MRI-derived abdominal skeletal muscle (SM) volume and DXA-derived trunk lean mass (A) and total lean mass (B) at baseline across the whole cohort.



Supplementary Table 1. Absolute change in MRI metrics across the whole cohort and individuals whose weight remained stable, who lost weight or gained weight over the follow-up.

Change	All indivduals[1] N = 3,029	10-15% weight loss [1] N = 55	5-10% weight loss[1] N = 255	2-5% weight loss[1] N = 538	Controls with stable weight[1] N = 1,333	2-5% weight gain N = 577	5-10% weight gain[1] N = 227	10-15% weight gain[1] N = 44	ANOVA p-value across all groups[2]
Liver									
Liver fat (%)	0.09 (2.70)	-3.23 (4.49)	-1.76 (3.03)	-0.73 (2.05)	0.07 (2.08)	0.96 (2.41)	2.19 (3.33)	3.45 (4.84)	<0.001
Body composition									
VAT (L)	0.10 (0.64)	-1.51 (0.90)	-0.80 (0.56)	-0.27 (0.36)	0.14 (0.34)	0.52 (0.38)	0.90 (0.61)	1.22 (0.60)	<0.001
SAT (L)	0.06 (0.83)	-2.23 (0.76)	-1.13 (0.62)	-0.42 (0.44)	0.09 (0.41)	0.59 (0.46)	1.13 (0.71)	2.07 (0.77)	<0.001
Abdominal SM (L)	-0.08 (0.21)	-0.26 (0.31)	-0.19 (0.23)	-0.14 (0.20)	-0.07 (0.19)	-0.02 (0.19)	0.03 (0.21)	0.07 (0.23)	<0.001
SMI (L/m^2)	-0.02 (0.09)	-0.08 (0.11)	-0.05 (0.09)	-0.04 (0.09)	-0.02 (0.08)	-0.01 (0.09)	0.01 (0.10)	0.03 (0.11)	<0.001
MFI_{SM} (%)	0.39 (0.58)	-0.57 (0.84)	-0.08 (0.58)	0.18 (0.50)	0.40 (0.47)	0.61 (0.51)	0.86 (0.62)	1.14 (0.61)	<0.001
Psoas CSA (cm^2)	-0.29 (2.51)	-0.98 (2.87)	-0.87 (2.14)	-0.57 (2.64)	-0.31 (2.55)	0.05 (2.33)	0.39 (2.61)	0.17 (1.69)	<0.001

[1] Data are mean (SD).

[2] P-values calculated using ANOVA test.

Supplementary Table 2. Relative change in MRI metrics across individuals who were not on medications for hypertension and whose weight remained stable, who lost weight or gained weight over 3-year follow-up.

Change	All indivduals[1] N = 2,648	10-15% weight loss [1] N = 50	5-10% weight loss[1] N = 219	2-5% weight loss[1] N = 461	Controls with stable weight[1] N = 1,172	2-5% weight gain N = 502	5-10% weight gain[1] N = 200	10-15% weight gain[1] N = 44	ANOVA p-value across all groups[2]
Liver									
Liver fat (%)	0.13 (2.56)	-3.23 (4.57)	-1.55 (2.58)	-0.65 (1.92)	0.08 (1.95)	0.97 (2.37)	1.99 (3.03)	3.45 (4.84)	<0.001
Body composition									
VAT (%)	6 (24)	-38 (15)	-22 (13)	-7 (23)	5 (13)	21 (19)	32 (23)	58 (38)	<0.001
SAT (%)	2 (17)	-29 (12)	-17 (10)	-7 (9)	2 (13)	12 (14)	20 (23)	36 (21)	<0.001
Abdominal SM (%)	-1.3 (3.5)	-4.1 (4.7)	-2.9 (3.4)	-2.2 (3.3)	-1.2 (3.3)	-0.4 (3.2)	0.5 (3.8)	1.2 (3.9)	<0.001
SMI (%)	-1.0 (4.8)	-3.5 (5.3)	-2.6 (4.4)	-1.9 (4.7)	-1.0 (4.5)	-0.3 (5.3)	0.6 (5.0)	1.6 (5.0)	<0.001
MFI_{SM} (%)	2.5 (4.0)	-3.7 (6.7)	-0.5 (3.8)	1.0 (3.5)	2.6 (3.1)	4.1 (3.4)	5.6 (4.6)	7.6 (4.4)	<0.001
Psoas CSA (c%)	0 (17)	-4 (18)	-4 (14)	-2 (18)	-1 (16)	1 (16)	4 (17)	6 (35)	<0.001

[1] Data are mean (SD) for relative % change.

[2] *P* values calculated using ANOVA test.

Supplementary Table 3. Relative change in MRI metrics in males whose weight remained stable, who lost weight or gained weight over the follow-up period.

Change	All males N = 1,486	10-15% weight loss [1] N = 20	5-10% weight loss[1] N = 123	2-5% weight loss[1] N = 271	Controls with stable weight[1] N = 683	2-5% weight gain N = 280	5-10% weight gain[1] N = 98	10-15% weight gain[1] N = 11	ANOVA p-value across all groups[2]
Liver									
Liver fat (%)	7 (42)	-41 (22)	-25 (23)	-9 (23)	5 (29)	25 (43)	62 (71)	92 (74)	<0.001
Body composition									
VAT (%)	4 (24)	-42 (16)	-24 (13)	-8 (28)	4 (13)	19 (16)	34 (25)	51 (25)	<0.001
SAT (%)	2 (18)	-32 (10)	-17 (9)	-7 (11)	3 (16)	13 (15)	22 (14)	44 (22)	<0.001
Abdominal SM (%)	-1.3 (3.6)	-5.4 (5.5)	-3.3 (3.6)	-2.7 (3.2)	-1.2 (3.4)|	0.1 (3.2)	1.0 (3.5)	2.8 (4.4)	<0.001
SMI (%)	-1.1 (4.6)	-4.9 (5.2)	-2.9 (4.5)	-2.4 (4.1)	-1.1 (4.2)	0.2 (4.9)	0.9 (5.0)	3.3 (5.9)	<0.001
MFISM (%)	2.6 (4.1)	-4.8 (9.8)	-0.5 (3.7)	1.1 (3.6)	2.7 (3.4)	4.2 (3.6)	5.7 (3.5)	7.8 (4.8)	<0.001
Psoas CSA (%)	-2 (14)	-9 (19)	-5 (11)	-3 (15)	-2 (15)	1 (15)	3 (14)	1 (6)	<0.001

[1] Data are mean (SD) for relative % change.

[2] *P* values calculated using ANOVA test.

Supplementary Table 4. Relative change in MRI metrics in females whose weight remained stable, who lost weight or gained weight over the follow-up period.

Change	All females N = 1,543	10-15% weight loss [1] N = 35	5-10% weight loss[1] N = 132	2-5% weight loss[1] N = 267	Controls with stable weight[1] N = 650	2-5% weight gain N = 297	5-10% weight gain[1] N = 129	10-15% weight gain[1] N = 33	ANOVA p-value across all groups[2]
Liver									
Liver fat (%)	8 (37)	-29 (32)	-16 (26)	-5 (25)	6 (29)	20 (39)	38 (48)	65 (69)	<0.001
Body composition									
VAT (%)	7(23)	-35 (14)	-20 (13)	-6 (11)	6 (13)	21 (20)	30 (21)	60 (41)	<0.001
SAT (%)	2 (16)	-28 (13)	-17 (11)	-6 (7)	2 (7)	11 (11)	18 (27)	33 (20)	<0.001
Abdominal SM (%)	-1.4 (3.5)	-3.9 (4.4)	-2.9 (3.4)	-2.2 (3.3)	-1.3 (3.3)	-0.9 (3.1)	0.2 (3.9)	0.6 (3.5)	<0.001
SMI (%)	-1.1 (4.9)	-3.3 (5.5)	-2.5 (4.2)	-1.8 (4.9)	-1.1 (4.6)	-0.7 (5.4)	0.4 (4.8)	1.0 (4.6)	<0.001
MFISM (%)	2.6 (3.8)	-3.1 (3.5)	-0.4 (3.8)	1.2 (3.3)	2.6 (2.9)	3.9 (3.2)	5.6 (5.0)	7.5 (4.3)	<0.001
Psoas CSA (%)	1 (19)	-2 (15)	-4 (16)	0 (20)	0 (17)	2 (17)	4 (21)	7 (41)	<0.001

[1] Data are mean (SD) for relative % change.

[2] *P* values calculated using ANOVA test.

Supplementary Table 5. Expected changes per each % weight reduction in individuals with overweight or obesity based on linear regression analysis.

Change[1]	Slope Coefficient	P-value	R^2
Liver			
Liver fat	4.7	<0.001	0.27
Body composition			
VAT	3.2	<0.001	0.71
SAT	2.2	<0.001	0.71
Abdominal SM	0.31	<0.001	0.20
SMI	0.29	<0.001	0.09
Psoas CSA	0.43	<0.001	0.02

[1] Shown as % change in MRI measure for each % reduction in weight.

References

1 Nowak M, Nunez L, Hill CE, Pagliaro T, McGonnigle J, Niglas M *et al.* Automated volumetric MRI quantification of body fat and skeletal muscle in the UK Biobank: technical validation against expert manual segmentations and comparison with single-slice techniques. *Abdomimal Radiology (NY)* 2025; **in press**.

2 Triplett WT, Baligand C, Forbes SC, Willcocks RJ, Lott DJ, DeVos S *et al.* Chemical shift-based MRI to measure fat fractions in dystrophic skeletal muscle. *Magn Reson Med* 2014; **72**: 8–19.

3 de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. *Radiology* 2004; **230**: 652–659.

4 Barral JK, Gudmundson E, Stikov N, Etezadi-Amoli M, Stoica P, Nishimura DG. A robust methodology for in vivo T1 mapping. *Magn Reson Med* 2010; **64**: 1057–1067.

5 Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA *et al.* N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010; **29**: 1310–1320.

6 Wasserthal J, Breit H-C, Meyer MT, Pradella M, Hinck D, Sauter AW *et al.* TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiology: Artificial Intelligence* 2023; **5**: e230024.

7 Xu W, Fu Y-L, Zhu D. ResNet and its application to medical image processing: Research progress and challenges. *Comput Methods Programs Biomed* 2023; **240**: 107660.