

## V. SUPPLEMENTARY MATERIAL

The following pages provide additional detail on predictive performance, sparsification, calibration, and inference with the proposed approach. Unless otherwise specified, all listed results were acquired with the configuration that combines both *mean-variance* regression and ensembling. The individual targets are furthermore examined in detail and compared to alternative UK Biobank reference values. Code samples for a PyTorch implementation are available online.

### **GitHub repository:**

<https://github.com/tarolangner/mri-biometry>

### *A. Datasets and Predictive Performance*

The effective number of samples in the three datasets used for evaluation is listed in Supplementary Table I, taking into account missing reference values. The inference dataset  $D_{inf}$  furthermore contained 29,234 and the repeat imaging dataset  $D_{revisit}$  another 1,179 unique samples.

**Supplementary Table I**  
NUMBER OF SUBJECTS PER DATASET

Field ID	Target		Cross-validation	Testing	Artifacts
22407	Visceral Adipose Tissue	(VAT)	8,534	1,096	327
22408	Abdominal Subcutaneous Adipose Tissue	(SAT)	8,534	1,097	326
22415	Total Adipose Tissue	(TAT)	8,270	0	242
22416	Total Lean Tissue	(TLT)	8,270	0	242
22409	Total Thigh Muscle	(TTM)	8,478	1,038	284
22436	Liver Fat Fraction	(LFF)	8,474	1,061	323

\*UK Biobank Field IDs and number of available subjects with known reference values per target in cross-validation on dataset  $D_{cv}$ , testing on dataset  $D_{test}$ , and artifact dataset  $D_{art}$ .

Supplementary Table II lists additional evaluation metrics on all targets in cross-validation and testing. The results of all four configurations in cross-validation are listed in Supplementary Table III.

**Supplementary Table II**  
PREDICTED PERFORMANCE IN DETAIL

Target	unit	Cross-validation					Testing					
		N	ICC	R <sup>2</sup>	MAE	MAPE	N	ICC	R <sup>2</sup>	MAE	MAPE	
Visceral Adipose Tissue	(VAT)	L	8,534	0.997	0.994	0.122	4.2	1,096	0.997	0.995	0.119	3.6
Abdominal Subcutaneous Adipose Tissue	(SAT)	L	8,534	0.996	0.993	0.191	2.8	1,097	0.996	0.992	0.192	2.7
Total Adipose Tissue	(TAT)	L	8,270	0.997	0.995	0.358	1.8	0				
Total Lean Tissue	(TLT)	L	8,270	0.983	0.966	0.579	2.5	0				
Total Thigh Muscle	(TTM)	L	8,478	0.996	0.993	0.162	1.6	1,038	0.995	0.990	0.174	1.6
Liver Fat Fraction	(LFF)	%	8,474	0.979	0.959	0.666	25.7	1,061	0.982	0.965	0.647	21.6

\* Results for an ensemble of ten *mean-variance* networks in 10-fold cross-validation on dataset  $D_{cv}$  and testing on dataset  $D_{test}$ . N: Number of subjects, ICC: Intraclass correlation coefficient, R<sup>2</sup>: Coefficient of determination, MAE: Mean absolute error, MAPE: Mean absolute percentage error.

Supplementary Table III  
COMPARISON OF ALL CONFIGURATIONS IN CROSS-VALIDATION

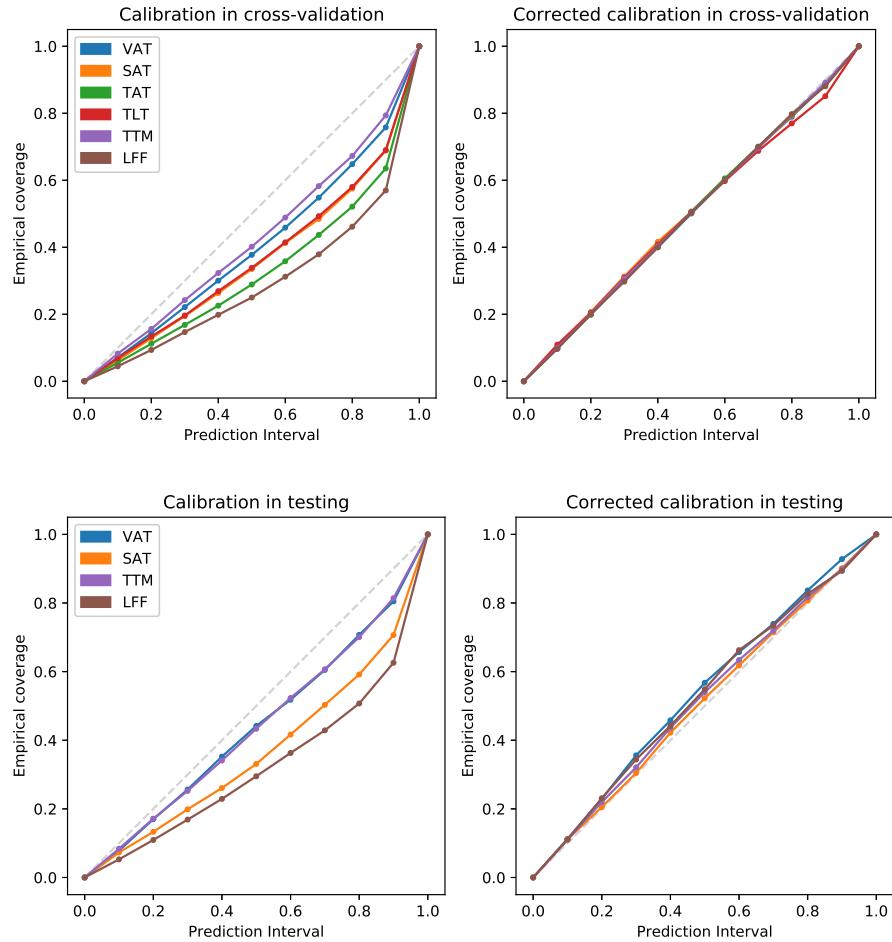
Configuration	ICC	R <sup>2</sup>	MAE	MAPE
<b>Visceral Adipose Tissue (VAT) in L</b>				
Least squares instance	0.996	0.992	0.150	5.2
Mean-variance instance	0.997	0.993	0.134	4.6
Least squares ensemble	0.997	0.993	0.133	4.6
Mean-variance ensemble	0.997	0.994	0.122	4.2
<b>Abdominal Subcutaneous Adipose Tissue (SAT) in L</b>				
Least squares instance	0.995	0.991	0.222	3.3
Mean-variance instance	0.996	0.992	0.209	3.1
Least squares ensemble	0.996	0.992	0.202	3.0
Mean-variance ensemble	0.996	0.993	0.191	2.8
<b>Total Adipose Tissue (TAT) in L</b>				
Least squares instance	0.997	0.993	0.420	2.1
Mean-variance instance	0.997	0.994	0.390	1.9
Least squares ensemble	0.997	0.994	0.377	1.9
Mean-variance ensemble	0.997	0.995	0.358	1.8
<b>Total Lean Tissue (TLT) in L</b>				
Least squares instance	0.981	0.963	0.650	2.8
Mean-variance instance	0.981	0.962	0.632	2.7
Least squares ensemble	0.983	0.966	0.594	2.6
Mean-variance ensemble	0.983	0.966	0.579	2.5
<b>Total Thigh Muscle (TTM) in L</b>				
Least squares instance	0.996	0.991	0.182	1.8
Mean-variance instance	0.996	0.992	0.176	1.8
Least squares ensemble	0.996	0.993	0.163	1.6
Mean-variance ensemble	0.996	0.993	0.162	1.6
<b>Liver Fat Fraction (LFF) in %</b>				
Least squares instance	0.977	0.956	0.706	28.1
Mean-variance instance	0.977	0.954	0.702	26.6
Least squares ensemble	0.979	0.960	0.671	27.0
Mean-variance ensemble	0.979	0.959	0.666	25.7

\* Results for all configurations in 10-fold cross-validation on dataset  $D_{cv}$ .

N: Number of subjects, ICC: Intraclass correlation coefficient, R<sup>2</sup>: Coefficient of determination, MAE: Mean absolute error, MAPE: Mean absolute percentage error.

### B. Overall Calibration

All examined configurations are biased towards overconfidence, consistently underestimating the true prediction errors. The predicted uncertainty should accordingly be scaled up. Suitable target-wise scaling factors can be determined to reach a better calibration on the validation data after training [28], [29]. In this work a simple grid search was used, which resulted in the target-wise scaling factors and the *areas under calibration error curve* (AUCE) [33] shown in Supplementary Table IV, with calibration plots, or reliability diagrams, shown in Supplementary Fig. 1. The same factors also achieve a considerable improvement when applied to the test data, indicating that the calibration of the proposed method could easily be corrected with this strategy for the normal material of the entire cohort.



Supplementary Figure 1. Calibration plots for the *mean-variance* regression ensemble on cross-validation dataset  $D_{cv}$  and testing on dataset  $D_{test}$ . Ideally, each prediction interval as modeled by the underlying predicted Gaussian probability distribution should cover the corresponding share of reference values. This hypothetical optimum is represented by the gray dashed line.

Supplementary Table IV  
CALIBRATION

Target		Cross-validation		Testing		Scaling factor
		AUCE	$AUCE_{scaled}$	AUCE	$AUCE_{scaled}$	
Visceral Adipose Tissue	(VAT)	0.089	0.003	0.052	0.035	1.90
Abdominal Subcutaneous Adipose Tissue	(SAT)	0.123	0.006	0.117	0.010	2.63
Total Adipose Tissue	(TAT)	0.154	0.003			3.35
Total Lean Tissue	(TLT)	0.120	0.012			2.50
Total Thigh Muscle	(TTM)	0.069	0.003	0.052	0.018	1.67
Liver Fat Fraction	(LFF)	0.186	0.003	0.156	0.027	4.34

\* Calibration of the *mean-variance* regression ensemble in cross-validation on dataset  $D_{cv}$  and testing on dataset  $D_{test}$ .

The area under calibration error curve (AUCE) can be far reduced (to  $AUCE_{scaled}$ ) with target-wise scaling factors.

### C. Detail on Individual Targets

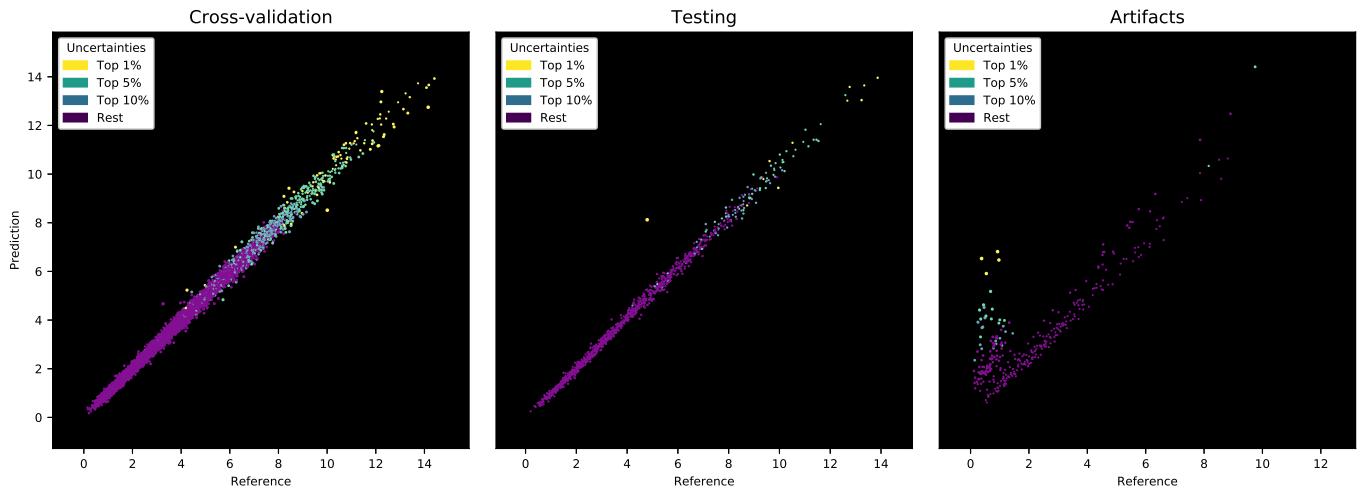
The following pages list dedicated plots for the prediction, sparsification, and calibration of each target. For the test data only the *mean-variance* ensemble configuration is shown, which was determined to be the best performing approach in cross-validation. Each subsection also includes short discussions and comparisons to alternative reference measurements which are primarily derived from two main sources. The first source contains body composition measurements obtained by Dual-energy X-ray absorptiometry (DXA) as conducted by UK Biobank [2]. The second source contains additional measurements based on independent machine learning analysis of the same neck-to-knee body MRI as used in this work as conducted by Application 23889, who have shared a return dataset 981.

Similar comparisons have been previously reported for a comparable *least squares* regression technique [17]. Some measurements may be highly correlated but yield low agreement due to a shift or scaling difference. Where specified, these alternative measurements were therefore mapped with linear regression to the target values as used in this work, so that agreement values can be reported. Additionally, Pearson's coefficient of correlation  $r$  is reported. For a fair comparison, the methods are evaluated on the same subjects.

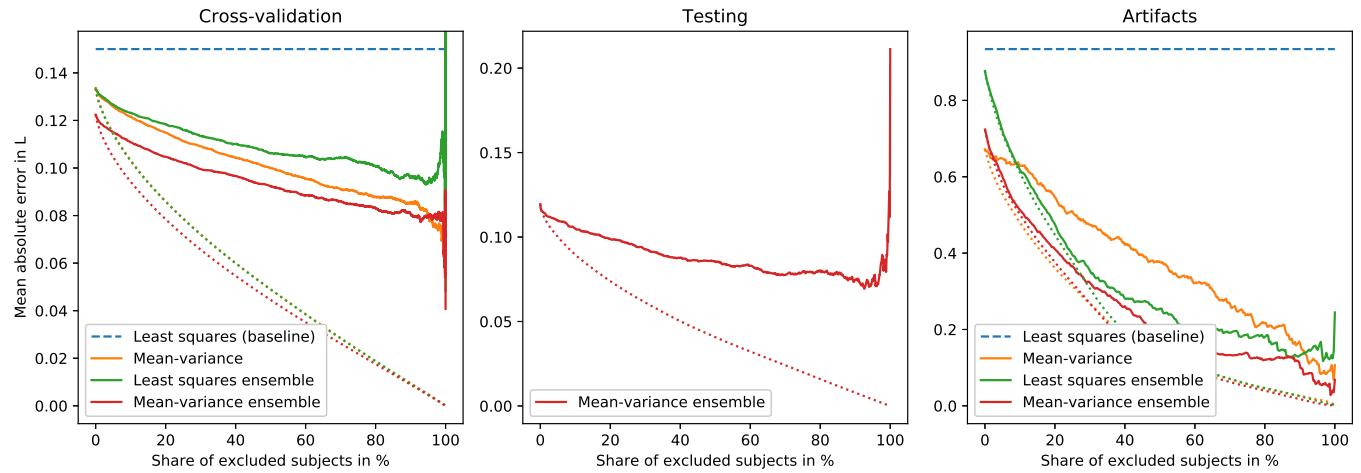
The sparsification plots also show *oracle sparsification curves* [26], which describe a hypothetical optimum that would result from sparsifying with a ranking of uncertainties that corresponds exactly to a ranking of absolute prediction errors. This optimum can typically not be reached in practice, as it would require imitating not only the desired measurements but also any inconsistencies and noise in the reference techniques themselves. The sparsification for the three evaluation datasets is shown separately, but it is worth noting that in most cases the samples with artifacts incurred the highest uncertainty. When applied to a dataset that included mixed normal material and artifacts, the latter would therefore typically be excluded first in the sparsification. The outlier with largest prediction error in testing for VAT, SAT, and TAT is the same subject, found to suffer from an atrophied right leg.

Aggregated saliency maps were obtained by generating guided gradient-weighted class activation maps for 3,091 subjects and co-aligning them by image registration [16]. Each aggregated saliency map accordingly highlights which anatomical structures were predominantly considered by the network to make predictions for the specified target. For clarity, the visualizations show the aggregated saliency as a heatmap for each of the three input image channels side by side and are provided with and without the template subject anatomy as an overlay. The network weights used for this purpose are based on the *mean-variance* configuration with a single network trained for cross-validation in this work, in each case using the instance that did not contain the given image in its training set.

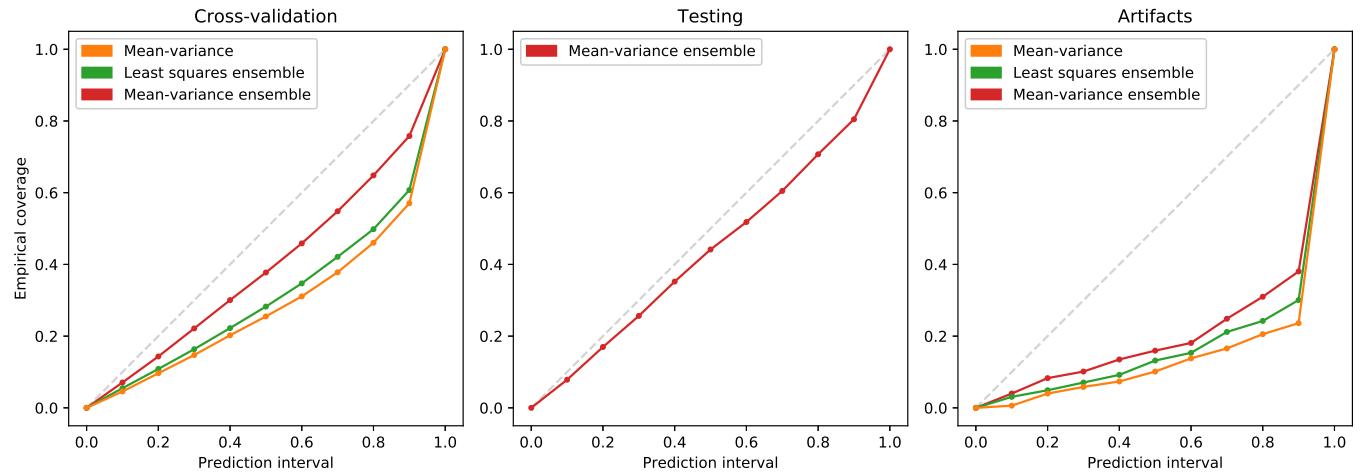
# I) : Visceral Adipose Tissue (VAT)



Supplementary Figure 2. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 3. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).



Supplementary Figure 4. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

### Visceral Adipose Tissue (VAT), extended notes:

Supplementary Fig. 2 shows a close fit with few outliers in the normal material. In testing, a single subject with an atrophied right leg incurs a substantially overestimated measurement, which can be identified by high uncertainty.

### Alternative reference methods:

UK Biobank field 23289 contains measurements of VAT by DXA for 5,109 subjects. These values were first converted from mL to L and then mapped to the target with the following linear transformation parameters:  $(2.27x + 0.83L)$ .

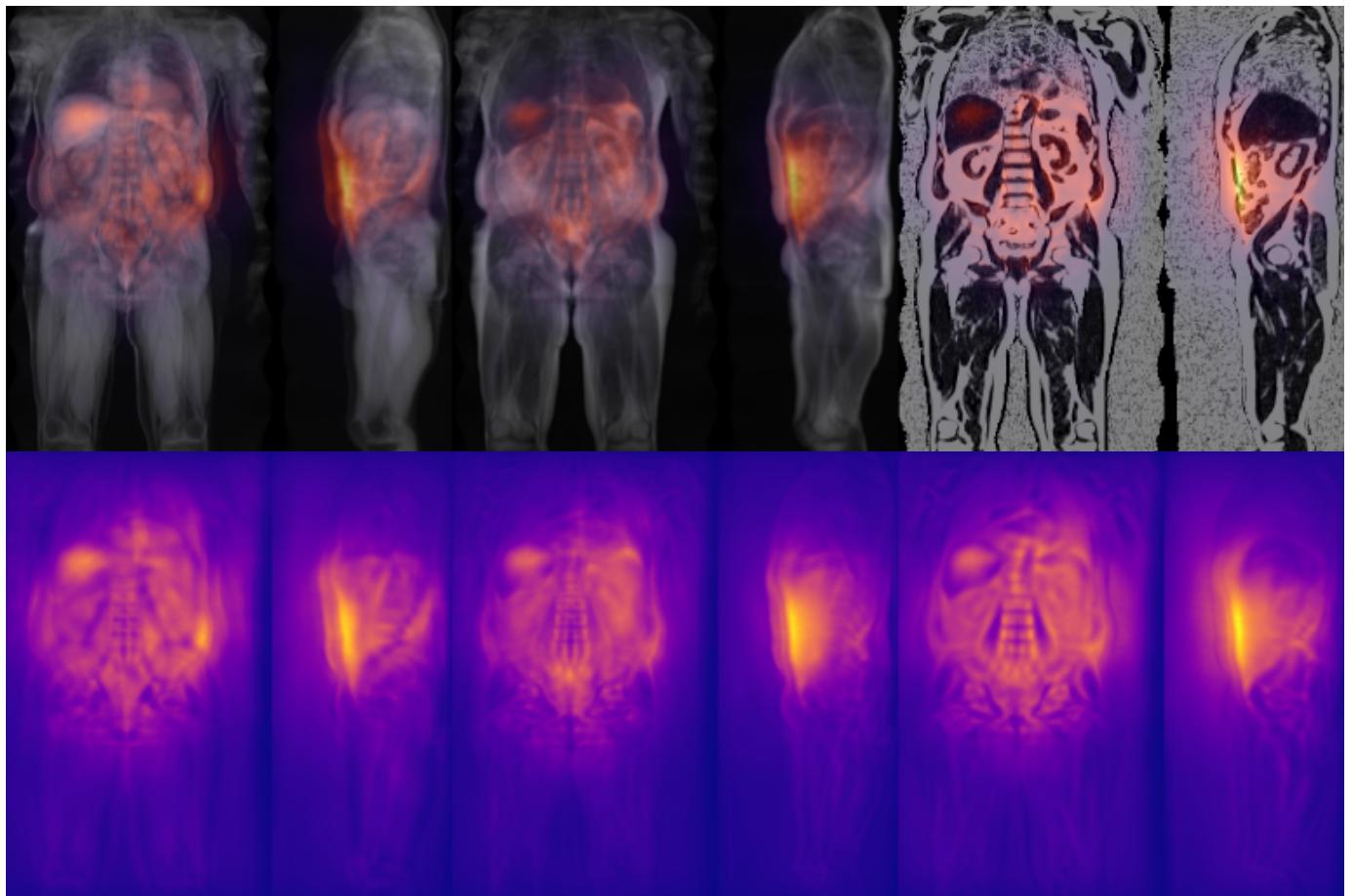
UK Biobank return 981 by application 23889 also offers VAT measurements for 9,127 subjects. These values were converted from mL to L, but did not require adjustment by linear regression.

Supplementary Table V  
COMPARISON OF VAT REFERENCES

Method	N	ICC	R2	MAE	MAPE	r
Proposed	4,491	<b>0.997</b>	<b>0.994</b>	<b>0.131</b>	<b>4.3</b>	<b>0.997</b>
Field 23289	4,491	0.970	0.942	0.401	14.9	0.971
Proposed	7,871	<b>0.997</b>	<b>0.994</b>	<b>0.121</b>	<b>4.1</b>	<b>0.997</b>
Return 981	7,871	0.996	0.993	0.137	4.4	0.996

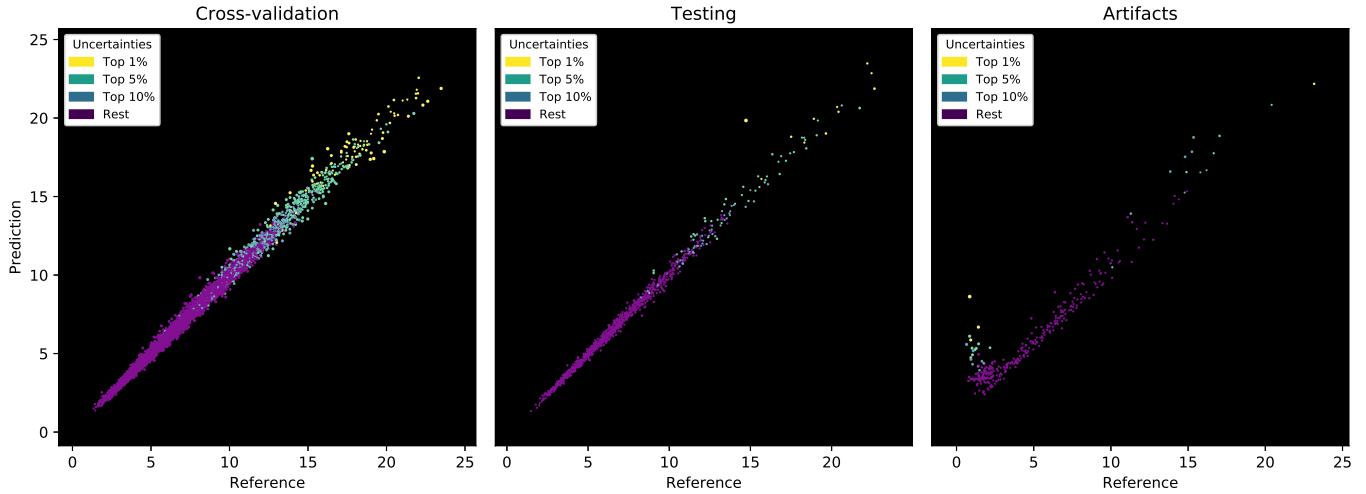
\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

### Aggregated saliency (VAT):

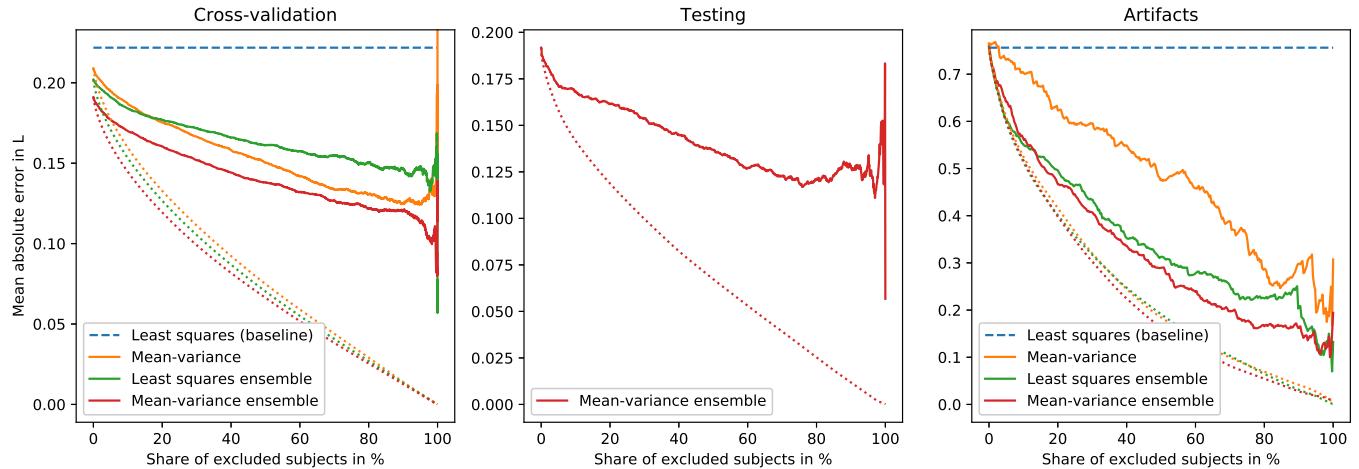


Supplementary Figure 5. Aggregated saliency [16] for Visceral Adipose Tissue (VAT) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

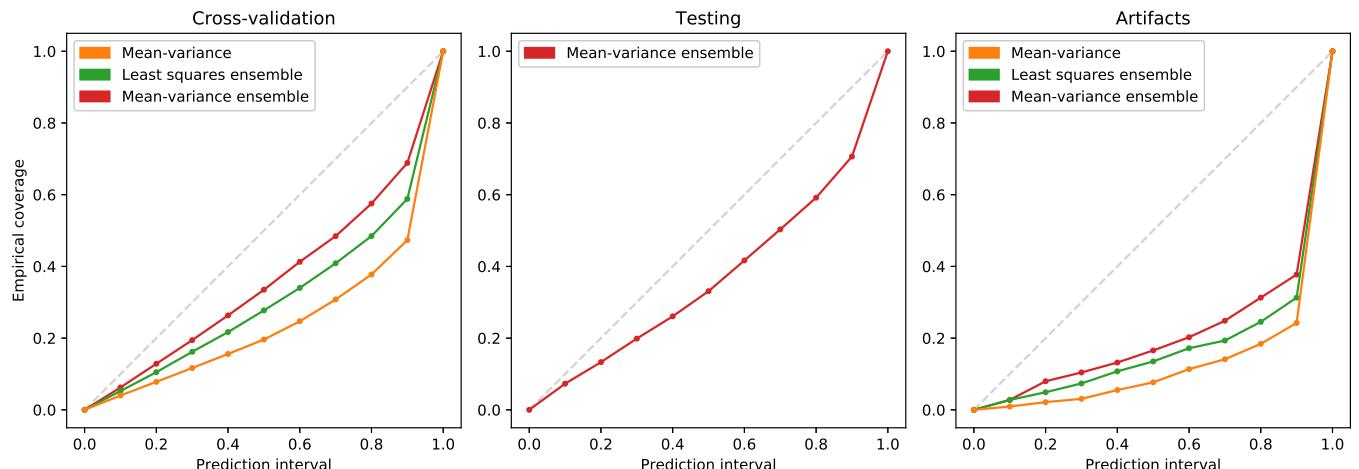
## 2) : Abdominal Subcutaneous Adipose Tissue (SAT)



Supplementary Figure 6. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 7. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).



Supplementary Figure 8. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

### Abdominal Subcutaneous Adipose Tissue (SAT), extended notes:

The scatter plot for the test data of Fig.6 shows a single outlier with about 15 L of subcutaneous adipose tissue, for whom the prediction yields almost 20 L with high uncertainty. This subject was found to suffer from an abnormal, atrophied right leg and also incurs high measurement errors in TTM and VAT.

### Alternative reference methods:

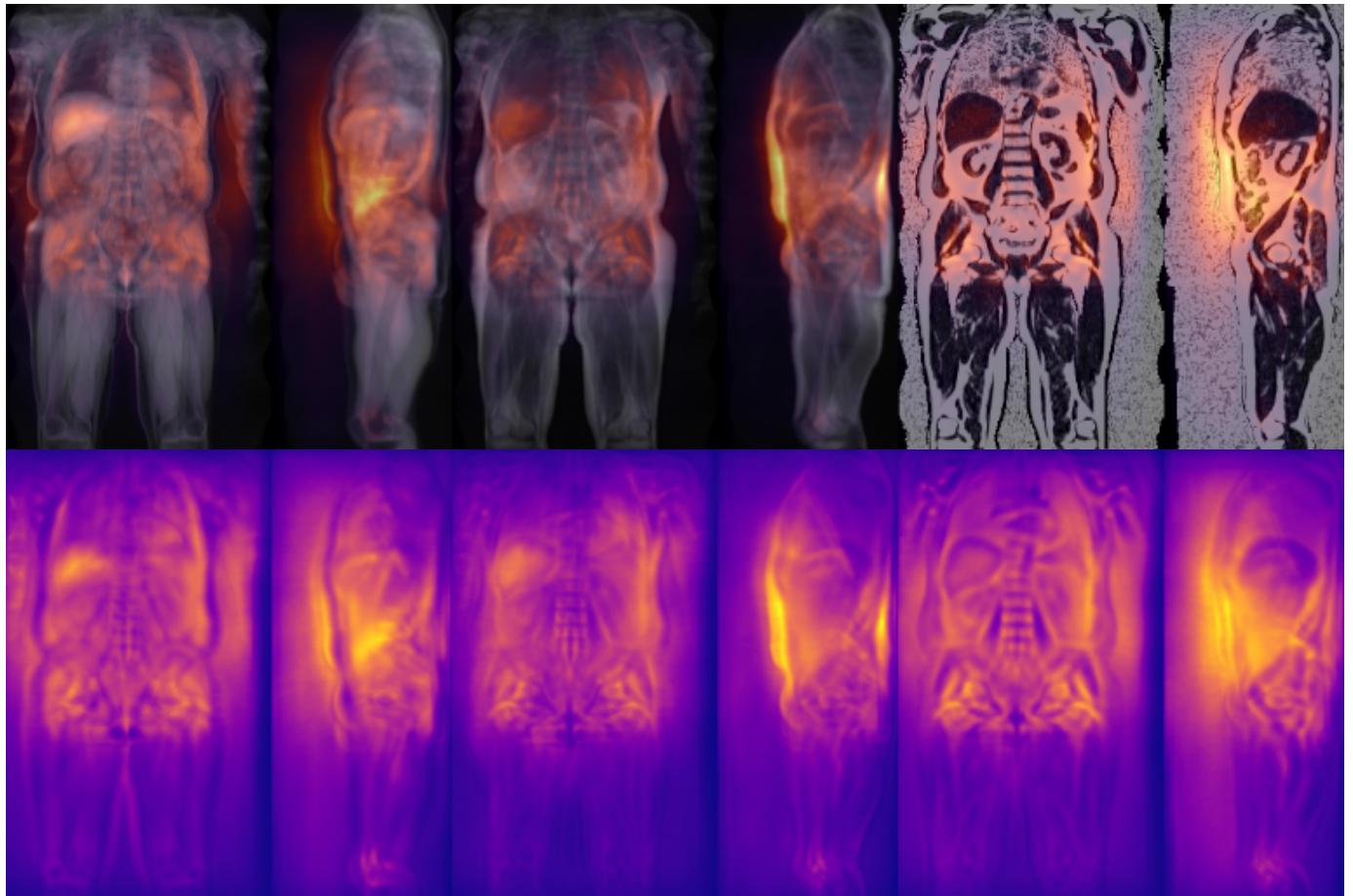
UK Biobank return 981 by application 23889 also offers measurements of subcutaneous adipose tissue volume for 9,379 subjects. These values were converted from mL to L and then mapped to the target with the following linear transformation parameters:  $(0.98x + 0.46L)$ .

Supplementary Table VI  
COMPARISON OF SAT REFERENCES

Method	N	ICC	R2	MAE	MAPE	r
Proposed	8,085	<b>0.996</b>	<b>0.993</b>	<b>0.187</b>	<b>2.8</b>	<b>0.996</b>
Return 981	8,085	0.994	0.989	0.208	3.1	0.994

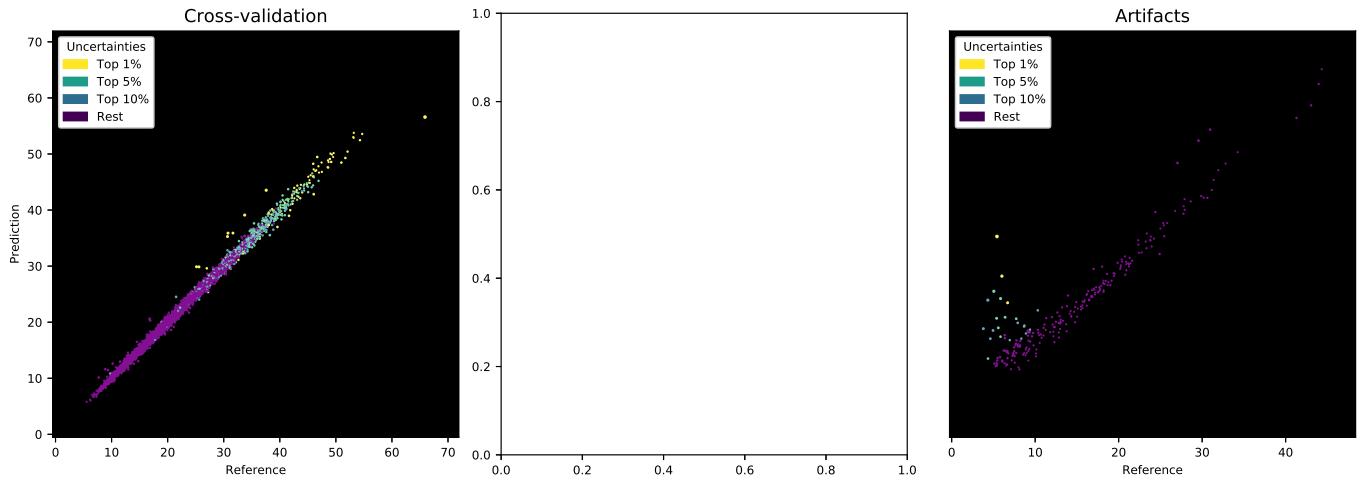
\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

### Aggregated saliency (SAT):

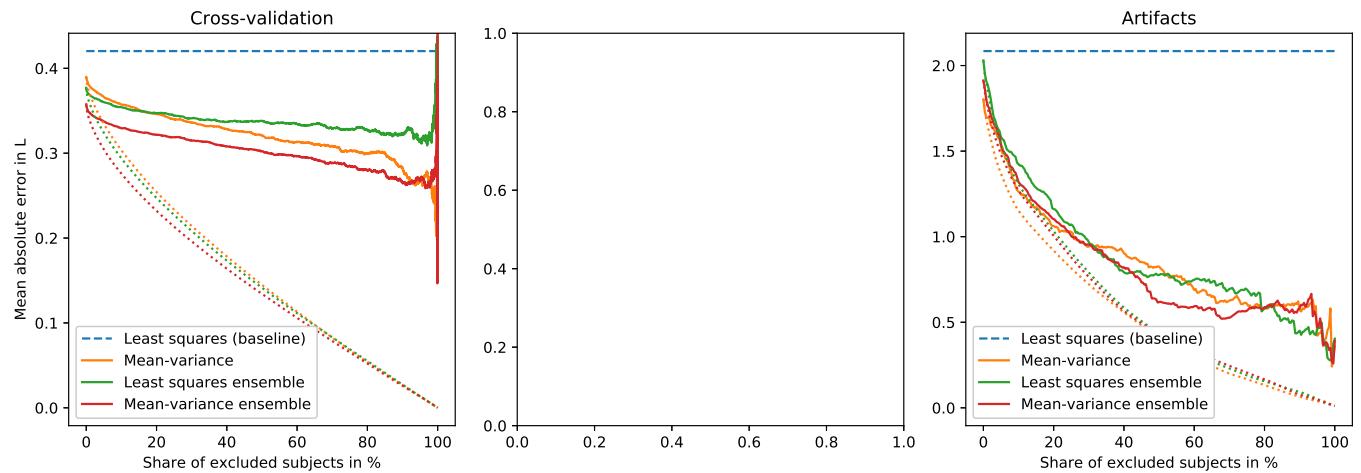


Supplementary Figure 9. Aggregated saliency [16] for Subcutaneous Adipose Tissue (SAT) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

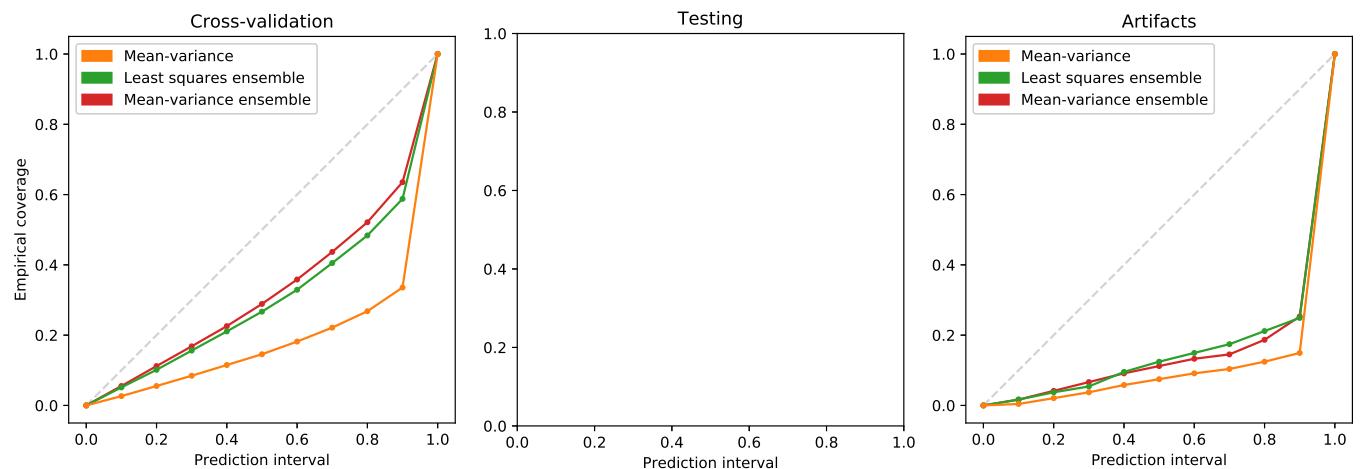
### 3) : Total Adipose Tissue (TAT)



Supplementary Figure 10. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 11. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).



Supplementary Figure 12. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

**Total Adipose Tissue (TAT), extended notes:**

No test data was available for this target.

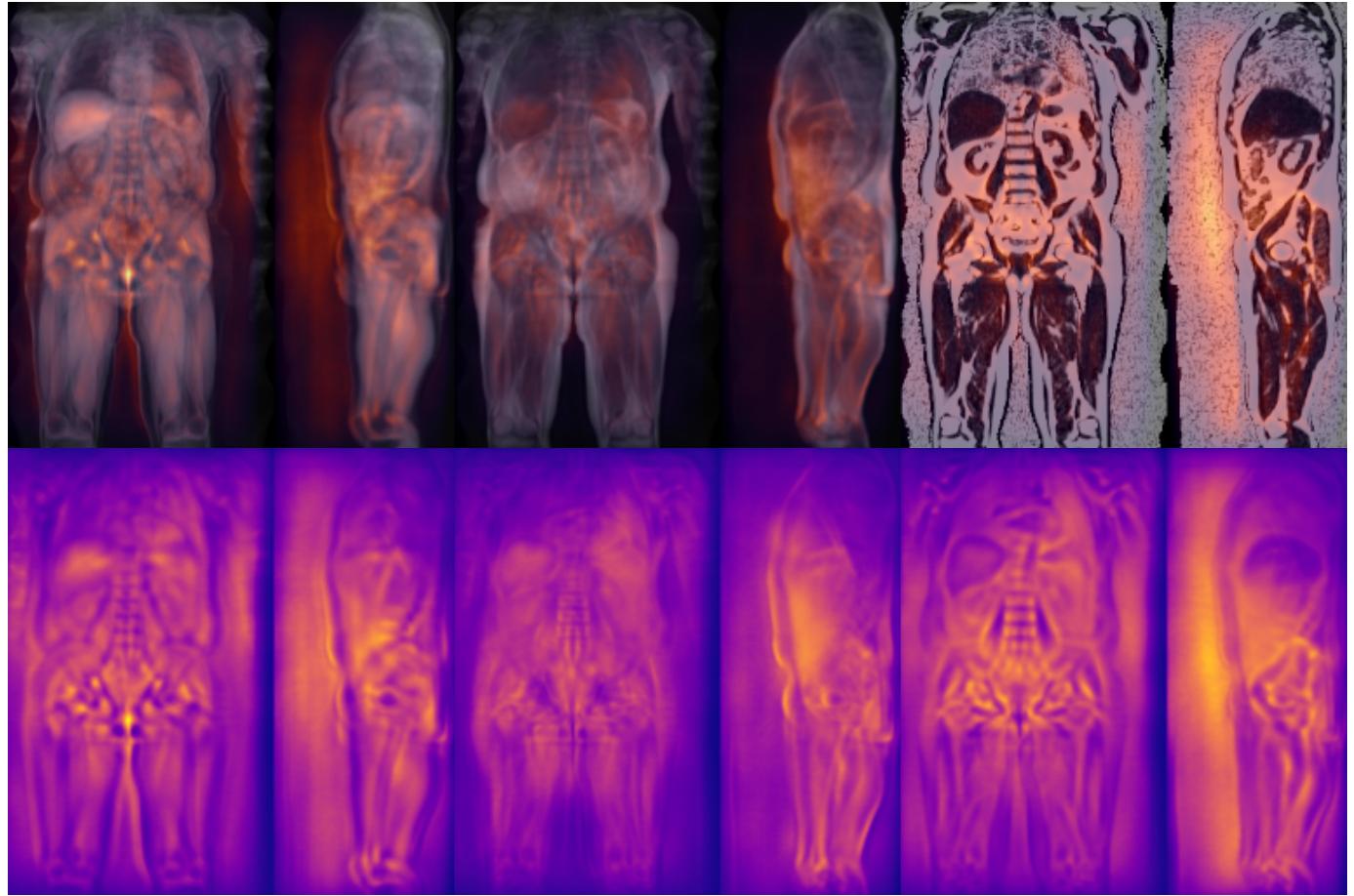
**Alternative reference methods:**

UK Biobank field 23278 contains alternative measurements of total fat mass by DXA for 5,170 subjects. These values were first converted from mL to L and then mapped to the target with the following linear transformation parameters:  $(0.80x + 0.51L)$ .

Supplementary Table VII  
COMPARISON OF TLT REFERENCES

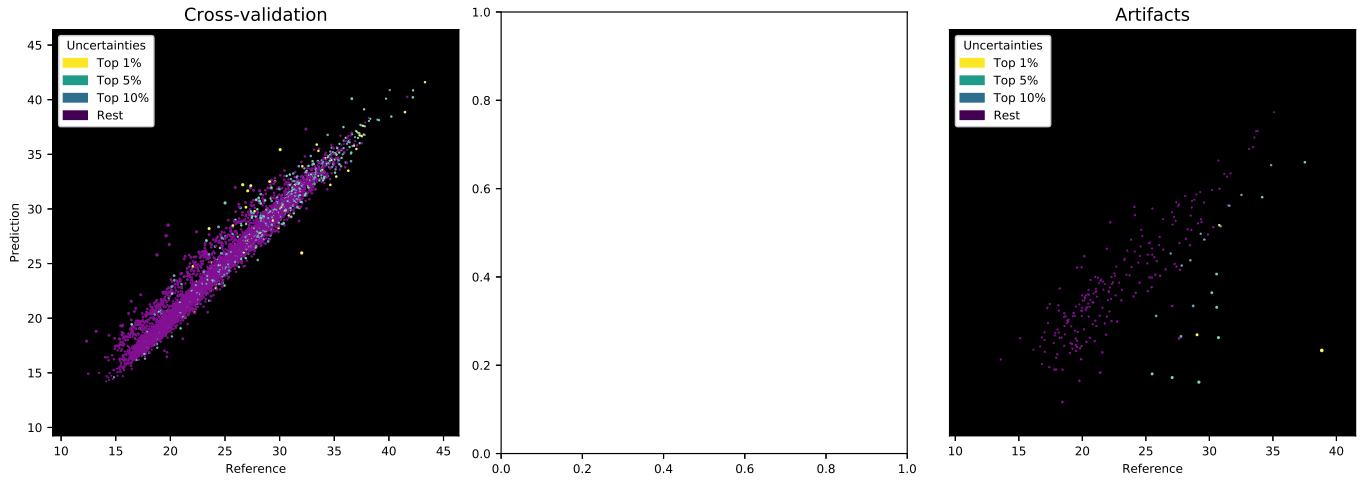
Method	N	ICC	R2	MAE	MAPE	r
Proposed	4,323	<b>0.997</b>	<b>0.995</b>	<b>0.353</b>	<b>1.8</b>	<b>0.997</b>
Field 23278	4,323	0.991	0.982	0.689	3.4	0.991

\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

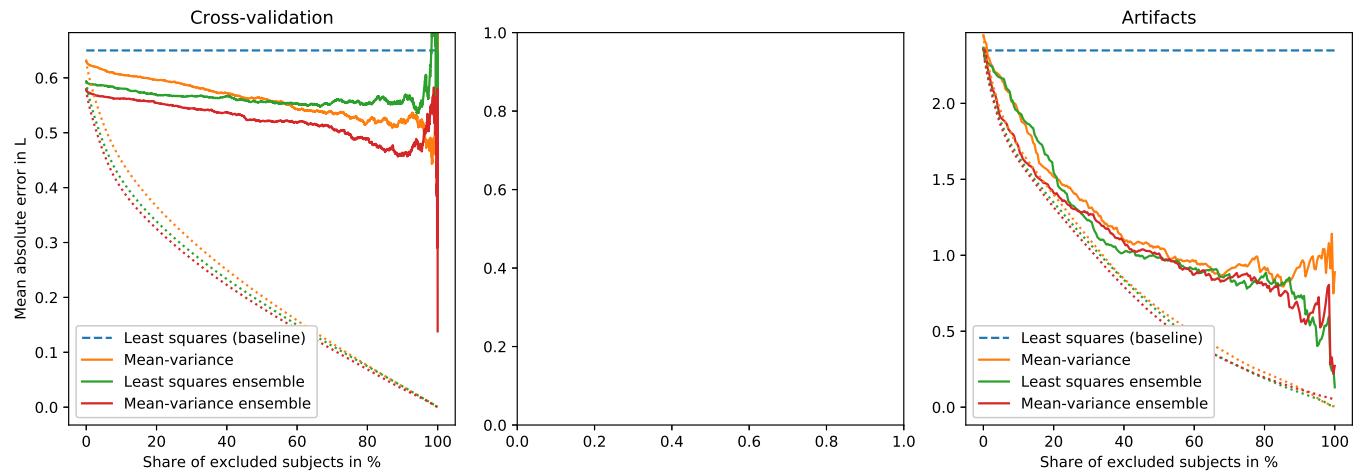
**Aggregated saliency (TAT):**

Supplementary Figure 13. Aggregated saliency [16] for Total Adipose Tissue (TAT) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

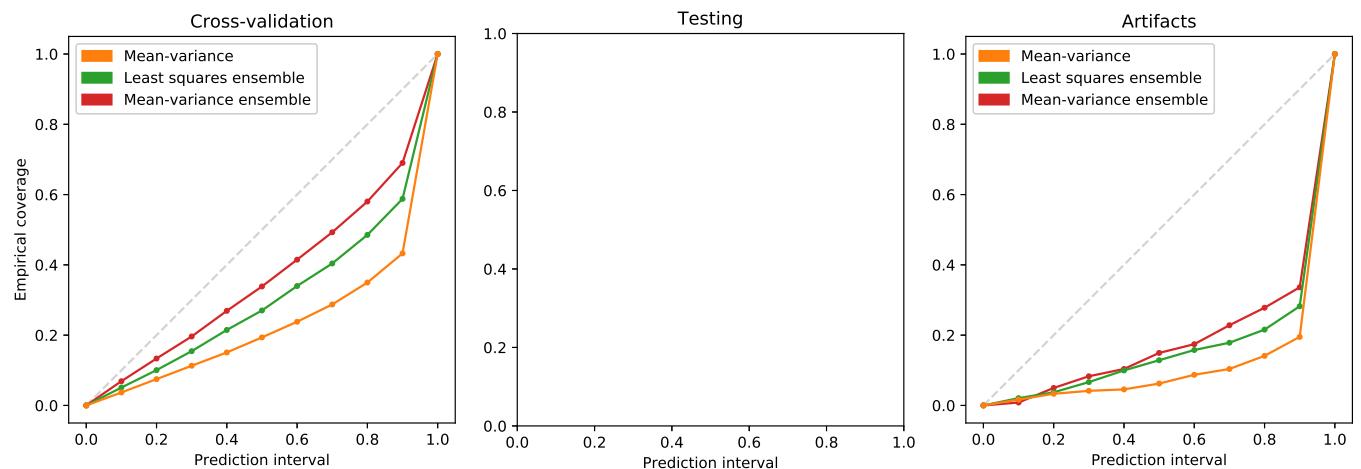
#### 4) : Total Lean Tissue (TLT)



Supplementary Figure 14. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 15. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).

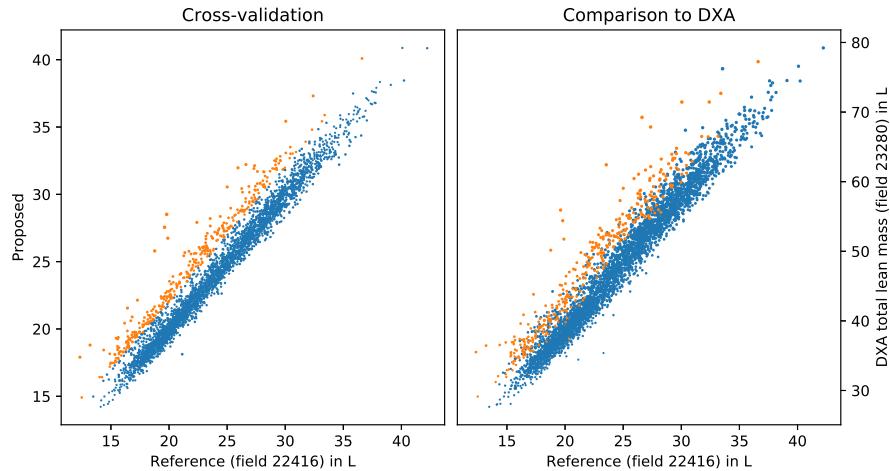


Supplementary Figure 16. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

### Total Lean Tissue (TLT), extended notes:

No test data was available for this target. Supplementary Fig. 14 shows a curious pattern for the cross-validation, where a subset of measurements is consistently overestimated by about 2 L.

The reason for this mismatch is unclear. The affected subjects are not part of the same cross-validation split set, were imaged in different imaging centers, and share no other obvious confounding factors. However, alternative measurements of total lean tissue by DXA (total lean mass, field 23280) independently support these overestimations relative to the reference used in this work. Supplementary Fig. 17 shows a comparison where the reference is plotted against the DXA measurements. All those cases that were overestimated by the proposed method by at least 2L are color-coded and form a similar pattern as observed in cross-validation.



Supplementary Figure 17. In some subjects (red), the proposed method overestimated total lean tissue (TLT) by at least 2L. As shown on the right, the DXA scan shows a similar pattern and independently indicates higher values for these subjects.

### Alternative reference methods:

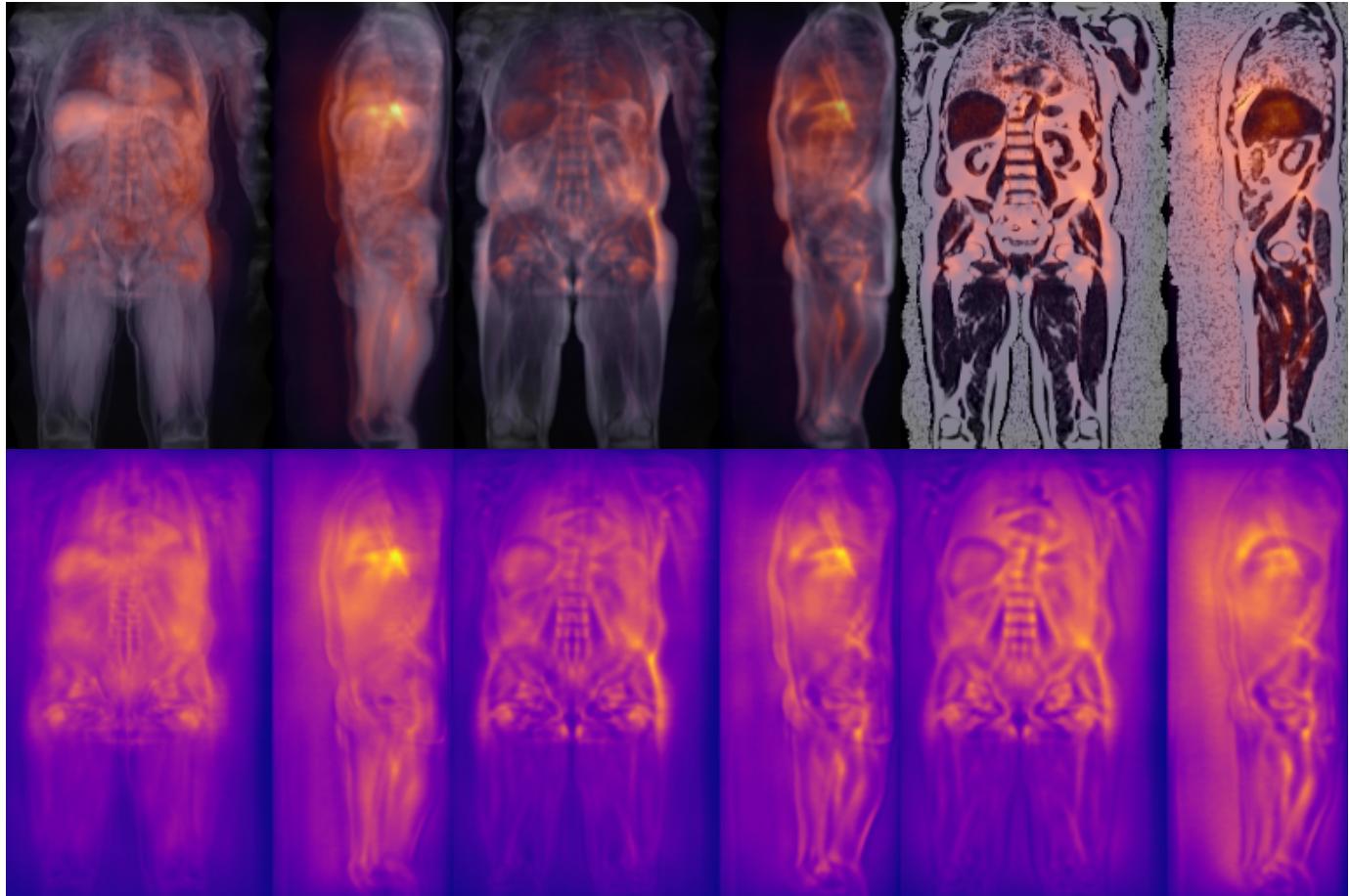
UK Biobank field 23280 contains additional measurements of total lean mass by DXA for 5,170 subjects. These values were first converted from mL to L and then mapped to the target with the following linear transformation parameters:  $(0.50x + 0.47L)$ .

On a side note, UK Biobank field 23285 also contains DXA measurements of trunk lean mass, but these values reaches lower agreement with the target than field 23280 and were not considered further.

Supplementary Table VIII  
COMPARISON OF TLT REFERENCES

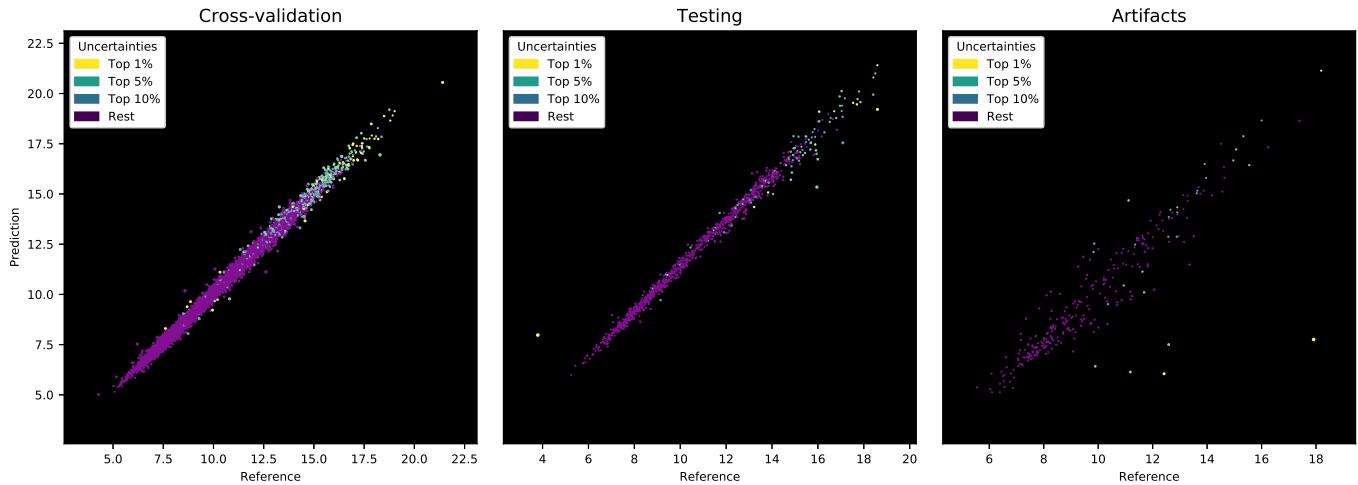
Method	N	ICC	R2	MAE	MAPE	r
Proposed	4,323	<b>0.976</b>	<b>0.953</b>	<b>0.684</b>	<b>3.0</b>	<b>0.978</b>
Field 23280	4,323	0.969	0.941	0.856	3.7	0.970

\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

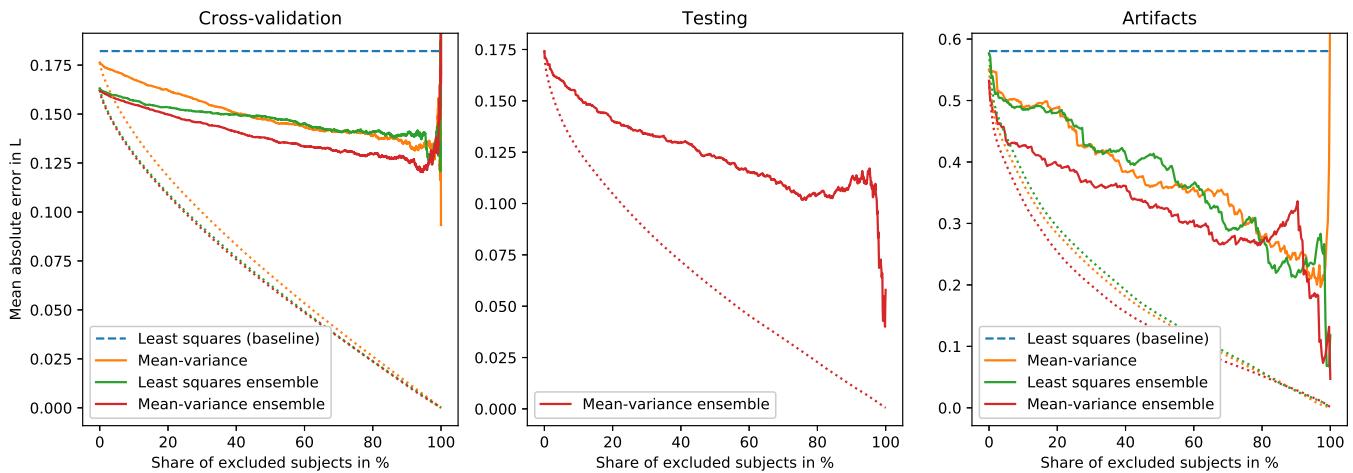
**Aggregated saliency (TLT):**

Supplementary Figure 18. Aggregated saliency [16] for Total Lean Tissue (TLT) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

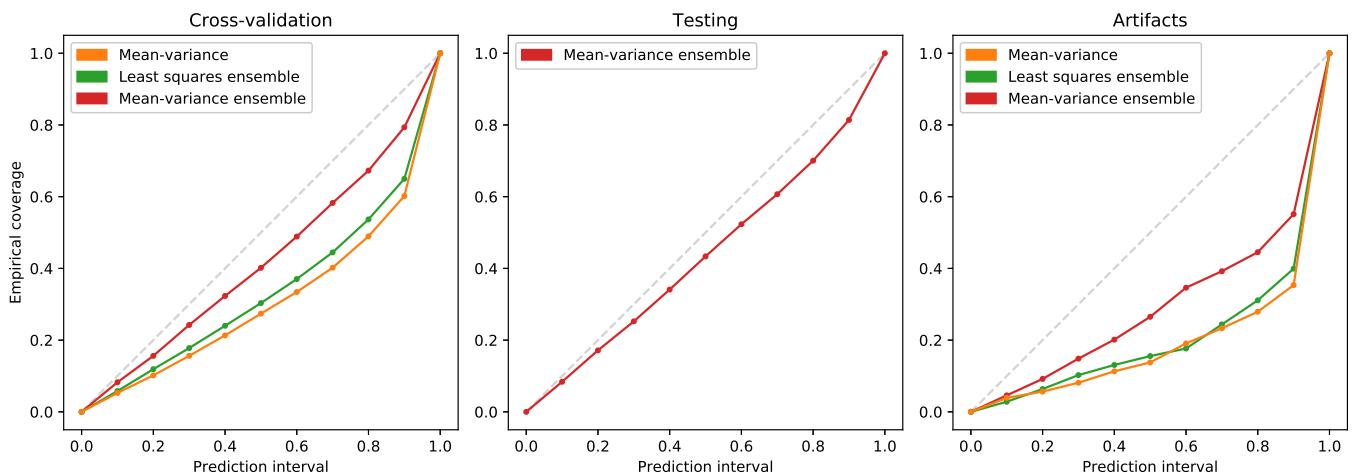
## 5) : Total Thigh Muscle (TTM)



Supplementary Figure 19. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 20. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).



Supplementary Figure 21. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

### Total Thigh Muscle (TTM), extended notes:

Supplementary Fig. 19 shows a close fit with few outliers in the normal material. In testing, a single subject with an atrophied right leg incurs high uncertainty, together with a moderately overestimated measurement. Several other high-valued testing cases are slightly underestimated. Many of those cases with the highest uncertainty show severe fat infiltrations of the thigh muscle.

### Alternative reference methods:

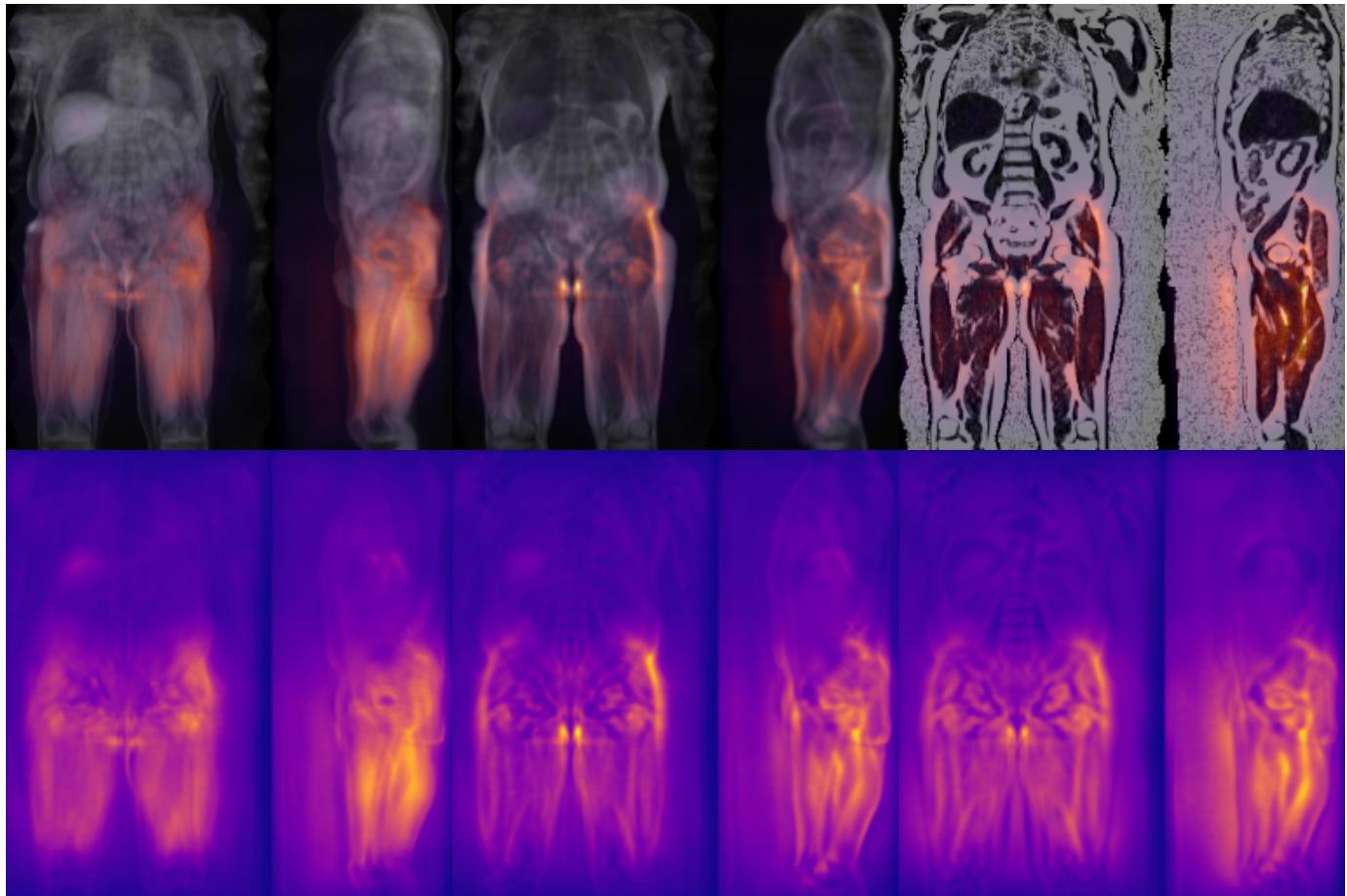
UK Biobank field 23275 contains measurements of the lean mass of the legs by DXA for 5,170 subjects. These values describe more than just muscle volume, but may still be considered as a proxy. These values were first converted from mL to L and then mapped to the target with the following linear transformation parameters:  
 $(0.69x + 0.64L)$ .

UK Biobank return 981 by application 23889 also offers thigh muscle volume measurements for 9,441 subjects. These values were first converted from mL to L and then mapped to the target with the following linear transformation parameters:  
 $(1.06x + 0.67L)$ .

Supplementary Table IX  
COMPARISON OF TTM REFERENCES

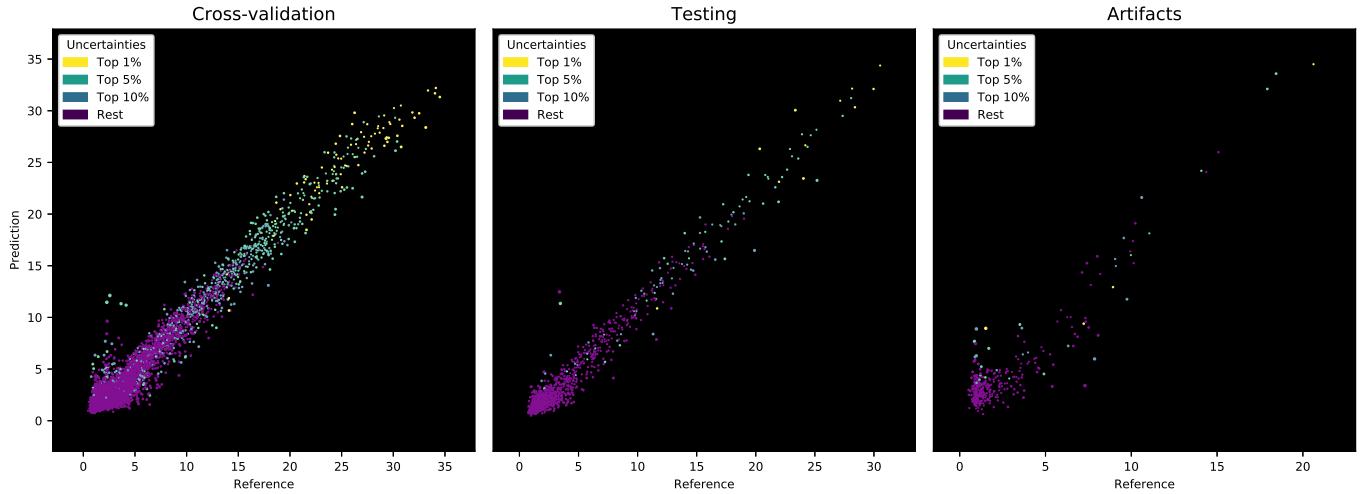
Method	N	ICC	R2	MAE	MAPE	r
Proposed	4,483	<b>0.996</b>	<b>0.992</b>	<b>0.173</b>	<b>1.7</b>	<b>0.997</b>
Field 23275	4,483	0.958	0.919	0.561	5.6	0.959
Proposed	8,144	<b>0.997</b>	<b>0.993</b>	<b>0.161</b>	<b>1.6</b>	<b>0.997</b>
Return 981	8,144	0.989	0.978	0.284	2.8	0.989

\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

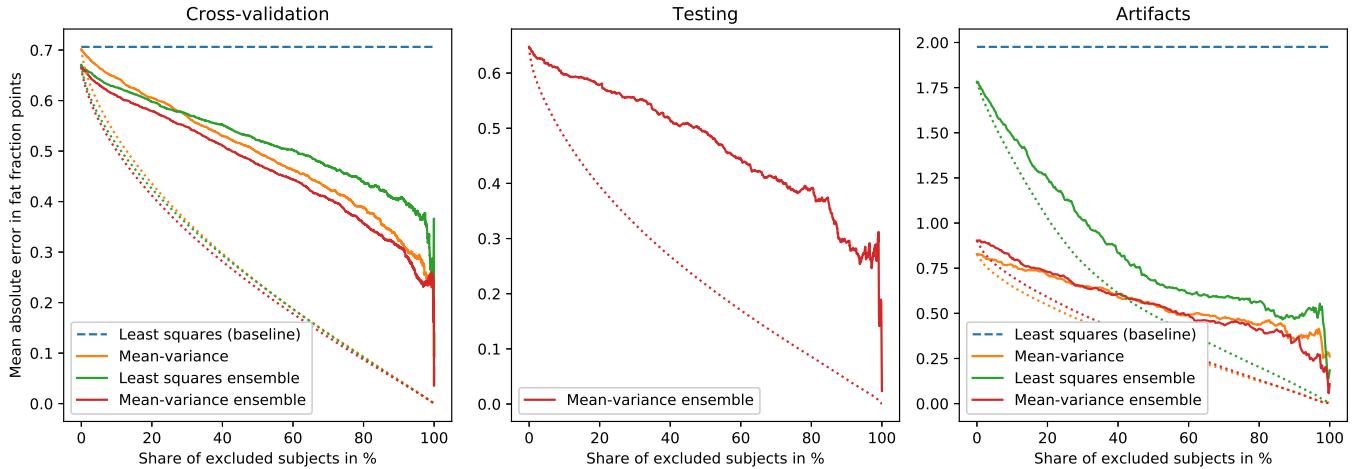
**Aggregated saliency (TTM):**

Supplementary Figure 22. Aggregated saliency [16] for Total Thigh Muscle (TTM) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

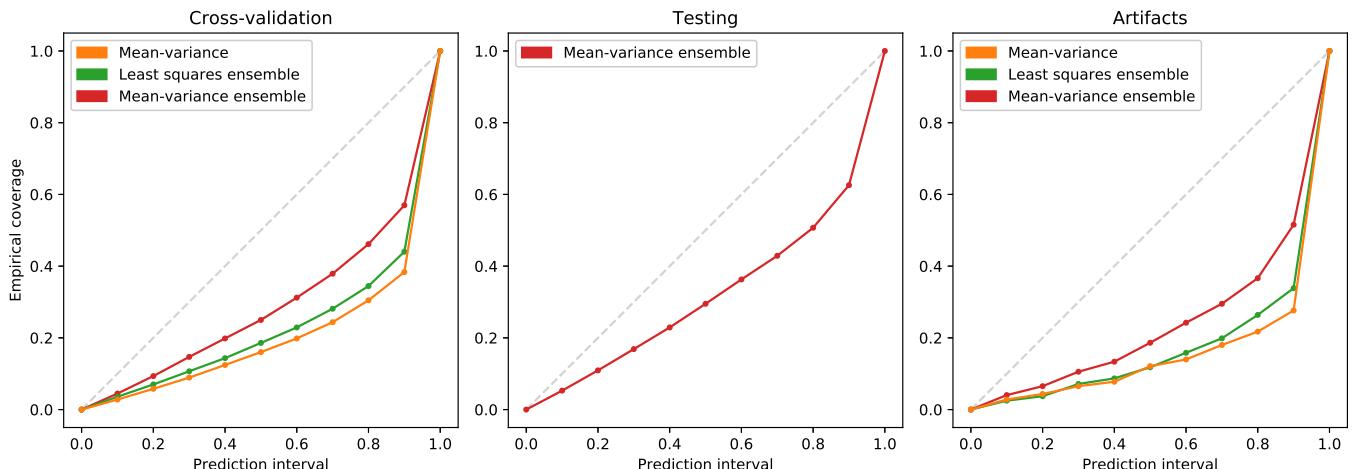
## 6) : Liver Fat Fraction (LFF)



Supplementary Figure 23. Predictions in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with color-coded uncertainty.



Supplementary Figure 24. Sparsification in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ , with oracle curves (dotted).



Supplementary Figure 25. Calibration in cross-validation on  $D_{cv}$ , testing on  $D_{test}$ , and on cases with artifacts of  $D_{art}$ .

### Liver Fat Fraction (LFF), extended notes:

The scatter plots of Supplementary Fig. 23 show that a small number of samples in the range of zero to five fat fraction points are severely overestimated, both in cross-validation and testing. Not all of these predictions incur high uncertainty.

Visual control of the affected subjects showed that the predictions by the proposed method often provided a better match to the neck-to-knee body MRI than achieved by the reference values. No obvious confounding factors such as artifacts or high liver iron content were observed. A similar effect was noted in previous work [21] where a *least squares* regression technique was trained to emulate an alternative set of UK Biobank liver fat measurements, field 22402. As both of these reference fields are based on the dedicated liver MRI instead of the neck-to-knee body MRI used here, a possible explanation could be an unusually severe mismatch of both protocols for these subjects.

On average, LFF incurred by far the highest normalized uncertainties (calculated by dividing the predicted uncertainty by the predicted means) of all targets. Finally, it is worth noting that for this target superior results may be possible when using an input format that only shows a fat fraction slice of the upper body, as previously proposed [21], although no rigorous comparison was attempted in the scope of this work. The technique could also be applied directly to the dedicated liver MRI.

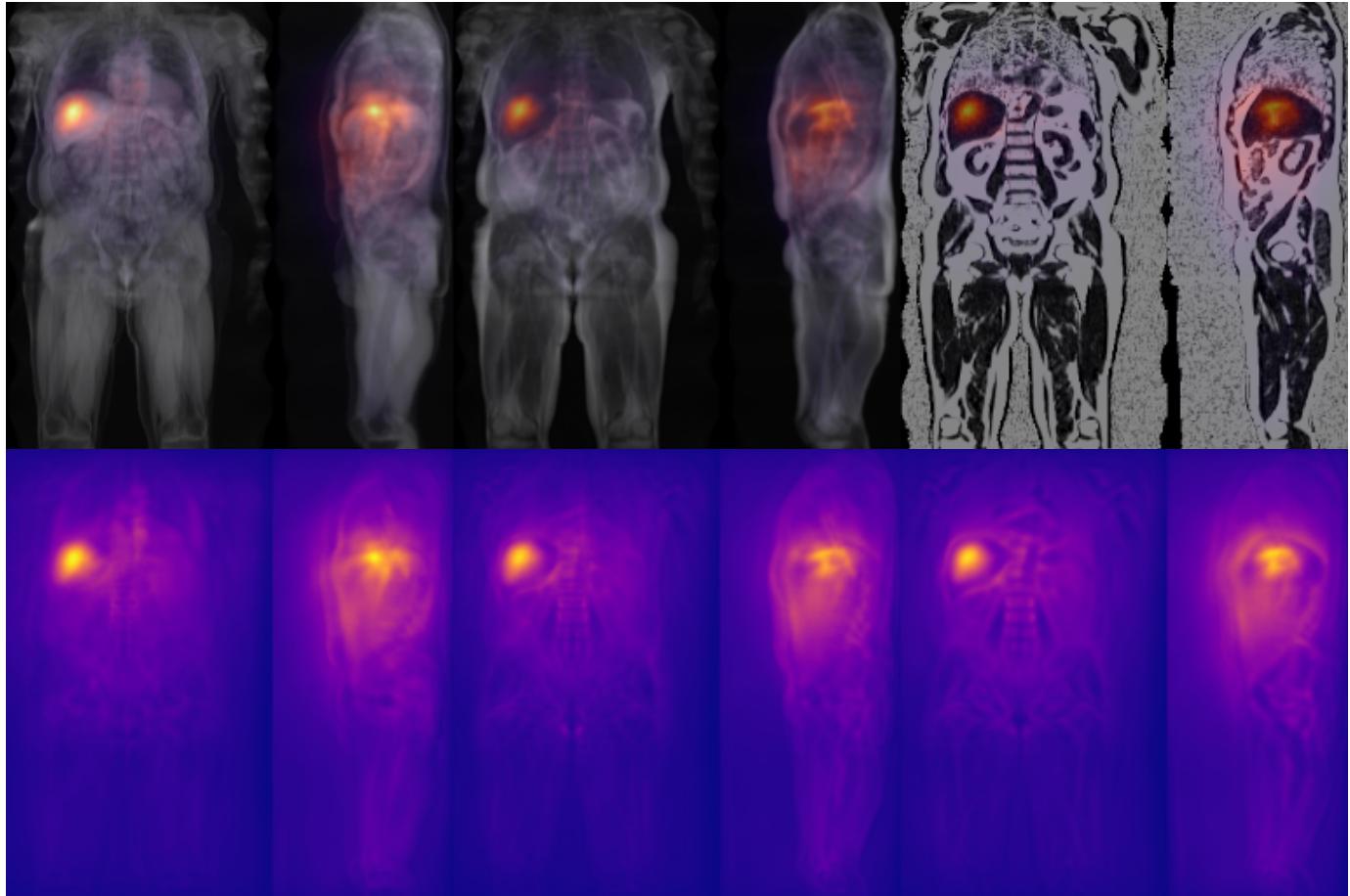
### Alternative reference methods:

UK Biobank field 22402 contains alternative liver fat fraction values for 4,616 subjects, obtained by mostly manual analysis of dedicated liver MRI [4]. Relative to the target used in this work, one outlier subject is overestimated by 24 fat fraction points and no linear transformation was applied.

Supplementary Table X  
COMPARISON OF LFF REFERENCES

Method	N	ICC	R2	MAE	MAPE	r
Proposed	4,401	0.978	0.956	0.669	26.3	0.978
Field 22402	4,401	<b>0.987</b>	<b>0.972</b>	<b>0.430</b>	<b>14.8</b>	<b>0.989</b>

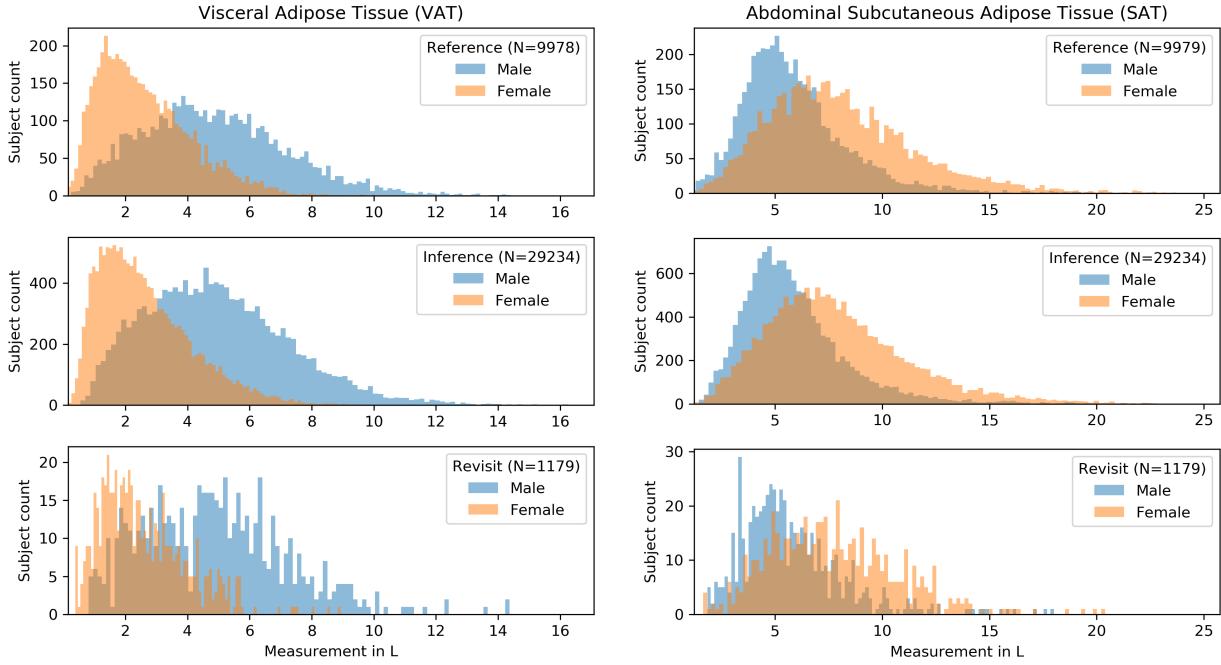
\*Comparison to the target values, listing both the proposed predictions and alternative UK Biobank reference values on the same subjects

**Aggregated saliency (LFF):**

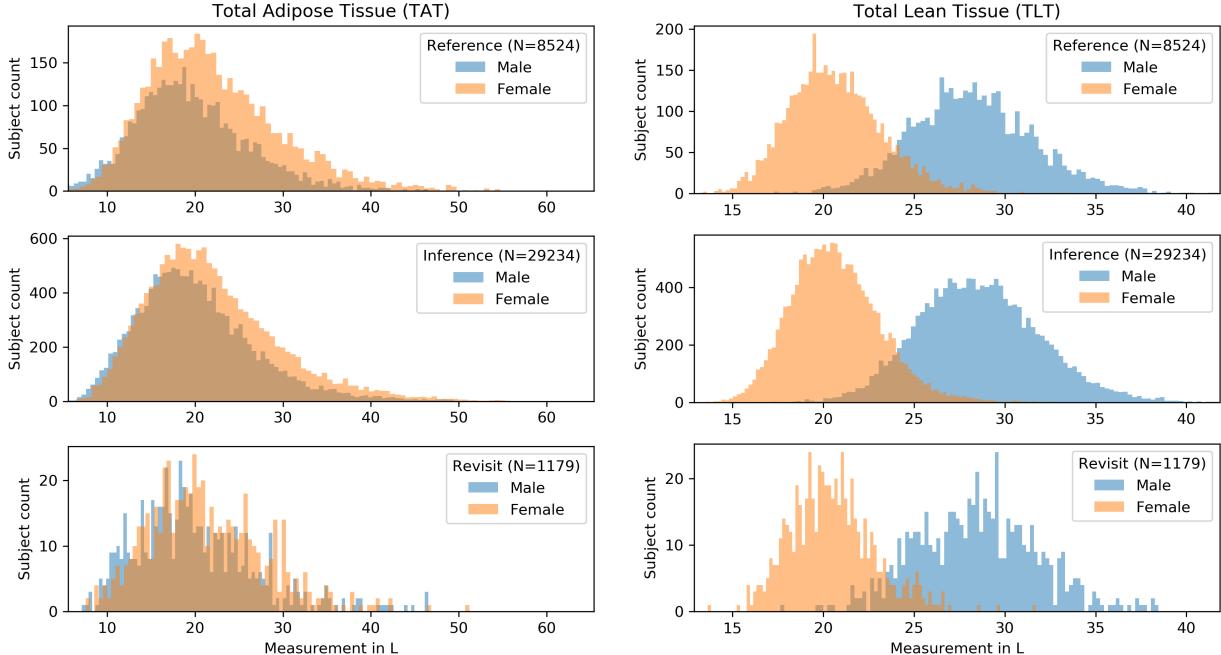
Supplementary Figure 26. Aggregated saliency [16] for Liver Fat Fraction (LFF) for 3,091 subjects, generated by a single *mean-variance* network. Each row shows the water, fat, and fat-fraction channels side by side, with the top row showing an overlay on the image data and the bottom row the saliency only.

#### D. Inference

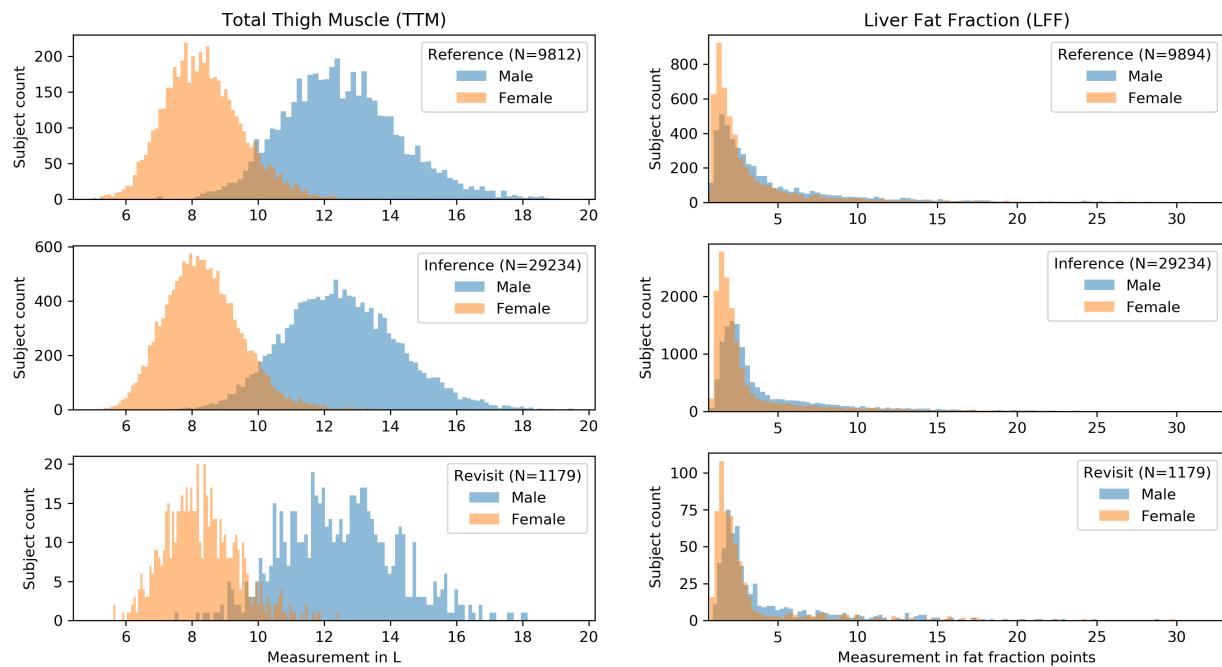
The following histograms of Supplementary Fig. 27, 28, and 29 show the reference values in comparison to those measurements predicted for inference on the original imaging visit on dataset  $D_{infer}$  and the later repeat imaging visit  $D_{revisit}$ . All shown data passed the visual quality controls, but no further attempt was made to exclude outliers based on the predicted uncertainty for these plots.



Supplementary Figure 27. Reference and predicted Visceral Adipose Tissue (VAT) (right column) and Subcutaneous Adipose Tissue (SAT) (right column).



Supplementary Figure 28. Reference and predicted Total Adipose Tissue (TAT) (left column) and Total Lean Tissue (TLT) (right column).



Supplementary Figure 29. Reference and predicted Total Thigh Muscle (TTM) (left column) and Liver Fat Fraction (LFF) (right column).