

# Prediction of Total and Regional Body Composition from 3D Body Shape

Chexuan Qiao<sup>\*1</sup>, Emanuella De Lucia Rolfe<sup>\*2</sup>, Ethan Mak<sup>1</sup>, Akash Sengupta<sup>1</sup>, Richard Powell<sup>2</sup>, Laura P. E. Watson<sup>3</sup>, Steven B. Heymsfield<sup>4</sup>, John Shepherd<sup>5</sup>, Nick Wareham<sup>2</sup>, Soren Brage<sup>\*2</sup>, and Roberto Cipolla<sup>\*1</sup>

<sup>1</sup>Department of Engineering, University of Cambridge, Cambridge, UK

<sup>2</sup>MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke’s Hospital, Cambridge, UK

<sup>3</sup>NIHR Cambridge Clinical Research Facility, Cambridge University Hospitals, Cambridge, UK

<sup>4</sup>Metabolism & Body Composition Laboratory, Pennington Biomedical Research Center, Baton Rouge, LA, USA

<sup>5</sup>Population Sciences in the Pacific Program (Cancer Epidemiology), University of Hawaii Cancer Center, Honolulu, HI, USA

## Abstract

Accurate assessment of body composition is essential for evaluating the risk of chronic disease. Access to direct imaging methods is limited due to practical and ethical constraints. 3D body shape correlates strongly with body composition. However, large-scale datasets containing 3D body shapes with paired anthropometric and metabolic traits are scarce. Here, we present a novel method that fits a 3D body mesh to a dual-energy X-ray absorptiometry (DXA) silhouette which is paired with anthropometric traits, using a large dataset from the UK population-based Fenland study (12,435 adults, age 30–65 years at baseline phase 1). We predict total and regional body composition metrics using these meshes, and monitor changes to body composition between the two Fenland phases. We also evaluate a 3D body shape smartphone app, which reconstructs a 3D body mesh from the phone images. We use these smartphone-generated meshes to predict body composition metrics, and compare the results against the reference methods DXA and air plethysmography (BODPOD). For the Fenland study, total and regional body composition metrics are predicted accurately, achieving  $r > 0.86$  for all metrics on phase 2 validation data. Accuracy is as high as  $r = 0.96$  for total fat mass, and  $r = 0.95$  for android fat mass. Predictions for changes achieve  $r > 0.60$  for all metrics. Using the smartphone-generated avatars, we achieve  $r > 0.84$  for all predicted metrics. We demonstrate that the 3D mesh approach is a valid alternative tool that can be easily implemented into clinical practice and large-scale population studies to track and monitor changes in body composition accurately.

## 1 Introduction

Body composition is strongly related to the risk of chronic disease morbidity and mortality [2], and can be assessed accurately using medical imaging methods such as dual-energy X-ray absorptiometry (DXA), magnetic resonance imaging (MRI) and computed tomography (CT) [6, 12, 44]. However, these methods are not readily available to be used routinely in clinical practice and in epidemiological studies due to practical and ethical constraints, nor are they easily accessible to the general public [10, 14]. In these settings, conventional anthropometry such as body mass index (BMI), waist, hip circumferences and waist-hip ratio are typically used to infer body composition. However, these indirect methods of assessing body composition

---

<sup>\*</sup>denotes equal contribution.

are insufficiently accurate or convenient for longitudinal use as they often require face-to-face clinical visits and trained staff. Furthermore, these surrogate measures do not differentiate between fat and lean mass or their distribution [31, 45].

There is a need to develop simple, accessible, and relatively inexpensive tools to improve the accuracy of assessing body composition. This would provide better prediction of metabolic health and identify people at high risk of disease in the long term so that remedial action can be taken. Significant works in recent years have focused on the development of 3D optical (3DO) scanning [43] to estimate body composition [3, 4, 29, 42]. 3DO scanners use depth sensors by projecting infrared patterns onto the scan subject to rapidly construct a 3D point cloud using multiview stereo, and subsequently capture 3D surface shape information. Rather than predicting body composition from anthropometric measurements alone, 3D body shape as a whole provides more visual and implicit cues for predicting body composition more accurately. Additional 3D shape cues can either be additional landmark diameters, circumferences, surface areas and volumes from 3DO scans [28], or parameters of a PCA shape space [3, 29, 42]. More recently, Leong et al. [24] use a variational autoencoder (VAE) [20] to learn latent DXA encoding, and map 3DO scans to pseudo-DXA images. These works have shown that 3D shape information could augment conventional prediction models using anthropometry, or outperform them as a standalone predictor for a variety of body composition metrics. However, while the cost of 3DO scanners is comparatively lower than that of DXA, MRI or CT, obtaining 3DO scans still requires a dedicated apparatus, which makes it less accessible to the general public.

To derive shape information without full reliance on 3DO scanners to reconstruct 3D body meshes, recent works have taken advantage of developments in computer vision and machine learning algorithms, which have enabled accurate segmentation [21] and pose estimation [36] of objects including the human body from RGB images, which are easily obtainable using a smartphone camera. Majmudar et al. [27] train a convolutional neural network (CNN) to directly predict percentage body fat from front and back images. Xie et al. [48] construct a PCA shape space from 2D DXA silhouettes and predict body composition. Tian et al. [41] derive a 3D body mesh from a pose-constrained 2D coronal silhouette and predict body composition. Sullivan et al. [40] derive body volume from a single image by measuring horizontal landmark diameters, and use the volume in a 3-compartment model (body mass, body volume and body water) to calculate percentage body fat. Smith et al. [38] compare circumference estimation accuracy using a smartphone app and 3DO scanners, and claim that circumferences estimated from images can be relatively accurate. McCarthy et al. [28] derive lengths, circumferences and volumes from body shape images, and predict skeletal muscle mass from these measurements alongside demographic variables. Graybeal et al. [11] evaluates two smartphone apps, and compare circumference and circumference ratio prediction accuracy. The remote data capture and modeling of 3D shape has numerous applications, including helping patients track individual changes over time for commonly assessed anthropometric measurements. Furthermore, patients are not required to physically attend clinics to have these measures done, thus lowering the burden on health services and providing a more cost-effective way to monitor aspects of patient health.

Unfortunately, there are few large-scale datasets containing 3D body meshes with paired anthropometric and metabolic traits. Bennett et al. [4] worked with a cohort of size 501, McCarthy et al. [28] had a cohort size of 322, Ng et al. [29] had a cohort size of 407. These small cohorts prevent larger deep learning models from being leveraged for body composition predictions. Klarqvist et al. [23] used a large-scale MRI database from the UK Biobank [7], to predict body composition from coronal and sagittal silhouettes. To the best of our knowledge, this is the only study that contains MRI data at this scale, as the use of MRI in population studies is limited due to cost and accessibility for research. 3D body shape datasets are therefore scarce, while datasets containing 2D DXA images with anthropometric and metabolic traits are in abundance. Therefore in this work, we present a novel method that first fits 3D body meshes to DXA silhouettes and paired anthropometry measurements consisting of height, waist and hip circumferences. We generate a large 3D database using our method. We then show that using the fitted meshes, total and regional body composition metrics can be predicted accurately.

We also test and evaluate the performance of a smartphone app (3D Body Shape App) that uses phone images alone to make it easier for individuals to visualise and track changes in their body shape [37]. The app captures four photographs (front, back, left-side, and right-side profiles of the participant), and reconstructs a 3D body mesh using these images. McCarthy et al. [28], Smith et al. [38] also generate 3D body meshes

from RGB images using an app, which is the most similar to our smartphone approach. However, their app requires constrained A-pose for the photographs, and could fail due to noisy background in our testing. In contrast, our method is robust to background, participant pose, camera orientation, and could be extended to accept an arbitrary number of input images. We show preliminary body composition prediction results using the app. Our aim is to improve performance by increasing reconstruction accuracy in the future.

In summary, we make the following contributions in this study:

- Construct a large 3D body shape database derived from 2D DXA silhouettes and paired anthropometry measurements (height, waist and hip circumferences);
- Predict, from 3D body shape, total and regional body composition metrics including:
  - Total fat mass;
  - Percentage body fat (PBF);
  - Android fat mass;
  - Gynoid fat mass;
  - Peripheral fat mass;
  - Visceral adipose tissue (VAT) mass;
  - Abdominal subcutaneous adipose tissue (SCAT) mass;
  - Total lean mass;
  - Appendicular lean mass;
  - Appendicular lean mass index (ALMI).
- Evaluate and show preliminary results of a smartphone app that predicts body composition by reconstructing 3D body meshes from images only.

To the best of our knowledge, our method is the first that fits 3D body meshes to DXA images, and predicts downstream body composition metrics. In this way we show that accurate 3D meshes can be derived from a single 2D silhouette plus simple anthropometry (height, waist and hip circumferences), from which body composition metrics can be predicted.

## 2 Methods

### 2.1 The Fenland Study

DXA scans, paired anthropometry data and metabolic health variables used for our method come from the Fenland study. Details of the study have been described elsewhere [25]. Briefly, the Fenland study is a population-based cohort study established in 2005. It comprises mainly participants of white European descent, born between 1950 and 1975, and recruited from general practice lists in Cambridgeshire (Cambridge, Ely, and Wisbech) in the United Kingdom. A total of 12,435 people took part in phase 1 of the study (2005–2015) and 7,795 in phase 2 of the study (2014–2020). Exclusion criteria for the Fenland study were pregnancy, diagnosed diabetes, inability to walk unaided, or psychosis. The study was approved by the Cambridge Local Research Ethics Committee and performed in accordance with the Declaration of Helsinki. All participants provided written informed consent to participate in the study.

In the current analyses, we excluded participants whose DXA scans had technical irregularities such as missing tissues or other scan artefacts. The analyses included 11,359 individuals (5,333 men and 6,026 women) from phase 1, 6,102 individuals from phase 2 (2,979 men and 3,123 women). 5,733 individuals participated in both phases. 80% of the Fenland phase 1 sample was used for the derivation and training of the 3D body shape composition models, and the remaining 20% of phase 1 was used to test the validity of those models. The phase 2 sample was used as a test dataset for validity in a now older population and to assess the sensitivity of prediction models to track within-individual changes over time.

## 2.2 Smartphone Validation Study

We also conducted a separate study which in addition to DXA scans also included air plethysmography (BODPOD) and a smartphone app capturing front, back and two side pose images. This was carried out in a sample of 119 healthy adults (39 men and 80 women) aged between 18 and 65 years old, free from disease and medications between July and November 2023. The participants in this study were recruited locally through advertisements and approval was granted by Cambridge central ethics committee (REC 06.Q0108.84). Written informed consent was obtained prior to the participants' visit. This independent sample was used to test the validity of our derived models from the Fenland study, as well as to evaluate the accuracy of the 3D shape obtained from smartphone images alone.

## 2.3 Anthropometry and Body Composition

In the Fenland study (both phase 1 and phase 2), weight was measured to the nearest 0.2kg with a calibrated electronic scale (TANITA model BC-418 MA; Tanita, Tokyo, Japan). Height was assessed to the nearest 0.1cm with a wall-mounted stadiometer (SECA 240; Seca, Birmingham, United Kingdom). Body mass index (BMI; in  $\text{kg}/\text{m}^2$ ) was calculated as weight divided by height squared. Waist circumference and hip circumference were measured to the nearest 0.1cm with a non-stretchable fiber-glass insertion tape (D loop tape; Chasmors Ltd, London, United Kingdom). Waist circumference was defined as the midpoint between the lowest rib margin and the iliac crest, and hip circumference was defined as the widest level over the trochanters. All measurements were taken by trained field workers. Body composition was assessed by DXA, a whole-body, low-intensity X-ray scan that precisely quantifies fat mass in different body regions (models used: Lunar Prodigy Advanced fan beam DXA scanner, or an iDXA system; GE Healthcare, Hatfield, UK). Participants were scanned supine by trained operators, using standard imaging and positioning protocols. All images were manually processed by one trained researcher, who corrected DXA demarcations according to a standardised procedure. In brief, the arm region included the arm and shoulder area (from the crease of the axilla and through the glenohumeral joint). The trunk region included the neck, chest, and abdominal and pelvic areas. The abdominal region (android region) was defined as the area between the ribs and the pelvis, and was enclosed by the trunk region. The leg region included all of the area below the lines that form the lower borders of the trunk. The gluteofemoral region (gynoid region) included the hips and upper thighs, and overlapped both leg and trunk regions. The upper demarcation of this region was below the top of the iliac crest at a distance of 1.5 times the abdominal height. DXA CoreScan software (GE Healthcare, Hatfield, UK) was used to determine visceral abdominal fat mass within the abdominal/android region. This software uses a proprietary inbuilt algorithm [1] to derive visceral abdominal fat mass within the android region, validated by the manufacturer using computed tomography and magnetic resonance imaging. The inbuilt algorithm estimates visceral abdominal fat mass by firstly estimating the subcutaneous fat width and the anteroposterior thickness of the abdominal wall. These parameters together with derived geometric constants are implemented to extrapolate the amount of subcutaneous fat mass in the android region. Visceral abdominal fat mass is then calculated by subtracting the estimated subcutaneous abdominal fat mass from the total android fat mass. Subcutaneous abdominal fat is therefore android fat mass minus visceral fat mass. The other body composition variables used in these analysis are derived as follows: Appendicular lean mass (ALM) is the sum of the lean tissue mass in the arms and legs. Appendicular lean mass is scaled to height to derive appendicular lean mass index ALMI ( $\text{ALM}/\text{height}^2$ ) [22]. Peripheral fat mass is the sum of the fat tissue mass in the arms and legs.

In the smartphone validation study, demographic information on age, sex, ethnicity was self-reported. Trained staff acquired all clinical measures. Height and weight were measured using a column scale (Seca GmbH & Co. KG, Hamburg). Waist and hip circumferences were measured using a tape measurer (CEFES-FIBRE by Hoechstmass Germany). Body volume was assessed using air plethysmography (BODPOD ADP system, Cosmed Srl, Rome, Italy), for which participants were in fitted clothing without shoes, and wearing a swim cap before entering the system. Total and regional body composition was measured using an iDXA scanner (GE Healthcare, Hatfield, UK). Four 2D photographs were captured by a smartphone camera (iPhone X, Apple Inc. IOS v15.6.1) using our purpose-built 3D Body Shape app, which constructs a 3D body mesh using phone images only. Participants wore form-fitting clothing, without shoes, and were asked to stand in an 'A' pose for the four photographs (front, back, left-side, and right-side profiles).

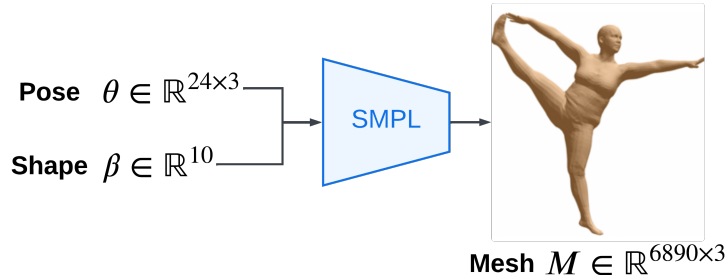


Figure 1: SMPL model which uses a low rank PCA representation of the body shape space. The model requires 24 joint rotation parameters and 10 shape parameters as input, and returns an expressive body mesh in a differentiable manner.

### 3 Model Derivation

Firstly, our method fits a 3D body mesh to a DXA silhouette with paired anthropometric measurements (participant height, waist and hip circumferences). Then, the fitted mesh shape parameters (SMPL shape  $\beta \in \mathbb{R}^{10}$  detailed in Section 3.1.1) are used to predict body composition metrics. Our smartphone validation study generates 3D body meshes from RGB images by averaging avatars across multiple views according to the uncertainty of shape parameters in each view [37].

This section describes our fitting pipeline, the body composition regressor, and smartphone avatar generation.

#### 3.1 Fitting the 3D Body Mesh to DXA Image

DXA images are single view and orthographic [33, 49], and hence lack depth information. To fit a 3D mesh, we therefore augment DXA silhouettes with paired anthropometry measurements, namely height, waist and hip circumferences.

Directly fitting a high dimensional point cloud to a single silhouette is challenging. We utilise Skinned Multi-Person Linear Model (SMPL) [26], a low rank PCA shape representation, which provides a strong prior for human body shape. Predicting SMPL pose and shape accurately in one go is also challenging, thus we utilise a two-stage approach where an optimiser refines an initial guess.

##### 3.1.1 SMPL

SMPL [26] is widely used for Human Pose and Shape (HPS) regression tasks [35–37]. Given input pose and shape parameters  $\theta \in \mathbb{R}^{24 \times 3}, \beta \in \mathbb{R}^{10}$ , SMPL returns a 3D mesh  $M \in \mathbb{R}^{6890 \times 3}$  in a fully differentiable manner. Fig. 1 shows the structure of the SMPL model.

##### 3.1.2 Initial SMPL Pose and Shape Prediction

To reconstruct a 3D mesh from a DXA image, we first make an initial guess of pose and shape using an off-the-shelf method. Most existing HPS networks take in RGB images as inputs and do not readily apply to DXA images [5, 17]. Instead, we utilise proxy representations of DXA images, consisting of edge images and joint heatmaps [35] in an attempt to bridge the domain gap between DXA and RGB images. Our initial pose and shape guess uses Hierarchical Kinematic Probability Distributions (HKPD) by Sengupta et al. [36], which adopts this proxy representation, and regresses probability distributions over SMPL pose and shape parameters. Regressing distributions also enables us to aggregate information across different views for the smartphone validation study detailed in Section 3.3.

While initial predictions from the HKPD method yield an estimation of body pose and shape from a single coronal view of the DXA participant, it is not sufficiently accurate for downstream metric regression tasks. The reason for this is two-fold: firstly, the coronal DEXA silhouette provides little information about the body shape in the depth dimension, which is important in terms of assessing body composition metrics such

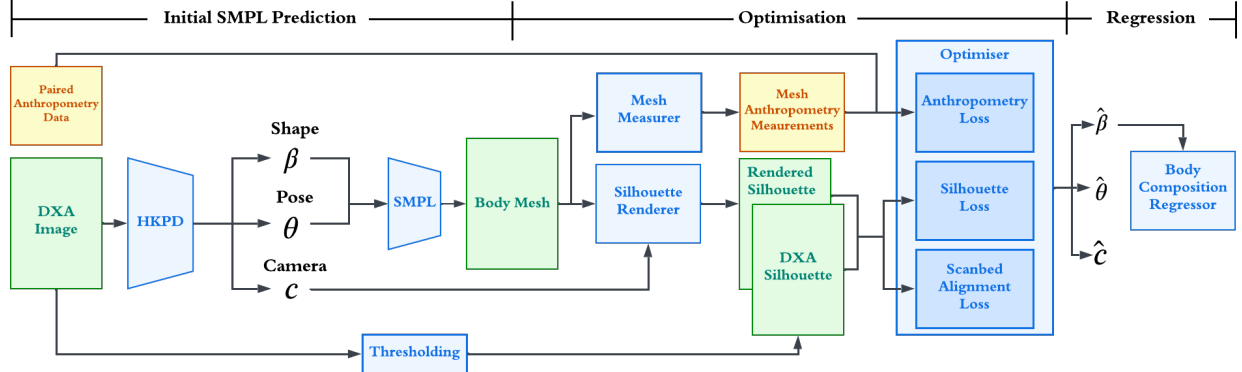


Figure 2: Our 3D body mesh fitting method. Given a DXA image, an initial pose and shape estimation is made using HKPD. Then, the output pose and shape parameters are optimised using losses constructed from DXA silhouettes and paired anthropometry data. Optimised shape parameters are used for body composition regression.

as visceral fat, which is correlated with the sagittal abdominal diameter (SAD) [15]; Secondly, HKPD is trained using synthetic body shapes sampled from Gaussian distributions, and tends to predict body shapes biased towards the SMPL mean.

### 3.1.3 Optimisation

Therefore, we construct an optimisation method to refine the initial guess, taking advantage of paired anthropometry measurements, DXA silhouettes, and further losses derived from special properties of DXA scanning (e.g. participants lying flat on the scanbed). We detail the optimisation losses below.

**Anthropometry loss.** We ensure that the mesh agrees with anthropometry measurements consisting of waist, hip circumferences and height by minimising the following loss:

$$\mathcal{L}_1 = \lambda_1 \|\hat{C}_W - C_W\|_2^2 + \lambda_2 \|\hat{C}_H - C_H\|_2^2 + \lambda_3 \|\hat{H} - H\|_2^2 \quad (1)$$

where  $\hat{\cdot}$  are measurements from the optimised mesh.  $[C_W, C_H, H]$  stand for waist, hip circumferences and height of the participant respectively. Height of the mesh is measured using the extrema vertices of the T-pose mesh.

To inject waist and hip information to the mesh whilst maintaining differentiability, we use a local ellipsoidal approximation to estimate circumferences. This is done by selecting a ring of keypoints around waist and hip, then fitting an ellipse using least squares. The selected keypoints are projected onto the horizontal plane, as SMPL vertices usually do not share the same height, and their relative heights vary across different body shapes.

**DXA silhouette loss.** Our optimiser also fits to the silhouette of the scan by differentially rendering the silhouette of the optimised mesh onto the image. We also impose a joint regulariser to prevent the optimised mesh from straying too far from the initial guess. This forms the second part of the loss function,

$$\mathcal{L}_2 = \lambda_4 \|\mathcal{R}(M(\hat{\theta}, \hat{\beta}), \hat{\mathbf{c}}) - S_{\text{gt}}\|_2^2 + \lambda_5 \|\hat{J}_{2\text{D}} - J_{2\text{D}}\|_2^2 \quad (2)$$

where  $\mathcal{R}(\cdot)$  is a PyTorch3D [32] differentiable silhouette renderer,  $\mathbf{c} = [s, t_x, t_y]$  are weak perspective camera parameters consisting of scale and translation.  $S_{\text{gt}}$  is the ground truth silhouette obtained by thresholding the DXA image.  $J_{2\text{D}}$  are 2D joint locations obtained by,

$$J_{2\text{D}} = s\Pi(\mathcal{J}M(\hat{\theta}, \hat{\beta})) + [t_x, t_y] \quad (3)$$

where  $\Pi(\cdot)$  is orthographic projection,  $\mathcal{J}$  is a linear vertex-to-joint regressor.

**Scanbed alignment loss.** Due to the orthographic nature of DXA images, pose of the DXA participant can be ambiguous from a single silhouette. As a result, initial guesses from HKPD often produce meshes that have forward-leaning torsos, or legs that are not fully extended. Since DXA participants are lying flat on the

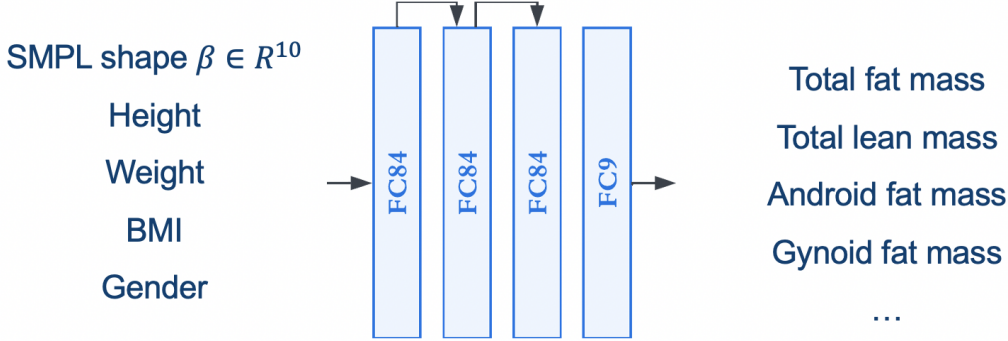


Figure 3: Our body composition regressor. The model is a simple feed-forward neural network which takes in optimised SMPL shape parameters, and outputs mass predictions.

scanbed, we impose an additional ‘scanbed alignment constraint’. During implementation we also impose a pose regulariser on the arms as well as a deviation from z loss to regularise arm poses.

$$\mathcal{L}_3 = \lambda_6 \|\hat{\theta}^{(1)}|_{x,y}\|_2^2 + \lambda_7 \|\hat{\theta}^{(2)} - \theta^{(2)}\|_2^2 + \lambda_8 \|\mathbf{z}_{\text{arms}} - z_{\text{pelvis}}\|_2^2 \quad (4)$$

where  $\theta^{(1)}$  is a subset of SMPL pose parameters containing spine and leg joints,  $\theta^{(2)}$  is a subset of SMPL pose parameters containing arm and wrist joints.

We optimise body pose, shape and camera parameters using the following total loss function,

$$\mathcal{L}(\beta, \theta, \mathbf{c}) = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \quad (5)$$

where  $\lambda_i$  are manually tuned to produce the best fits. Fig. 2 shows our full mesh fitting pipeline. Optimised SMPL shape parameters  $\hat{\beta} \in \mathbb{R}^{10}$  are used for the next phase which regresses body composition metrics.

### 3.2 Body Composition Prediction

We construct a simple feed-forward neural network to regress body composition metrics from the 3D meshes. The inputs to the network are the 10 SMPL shape parameters obtained from DXA optimisation, height, weight, BMI, and gender of the participant. The outputs of the network are the estimated body composition metrics such as total fat mass, total lean mass, etc. We use residual connections in the first two layers as they slightly improve our lean mass predictions. The structure of our network is shown in Fig. 3.

The network is trained using a mean squared error (MSE) loss function constructed from target and predicted mass values. We weight the loss function using homoscedastic uncertainty [18], and learn these metric-wise uncertainties automatically,

$$\mathcal{L}(\mathbf{W}, \sigma) = \sum_i \frac{1}{2\sigma_i^2} \|\hat{y}_i - y_i\|_2^2 + \sum_i \log \sigma_i \quad (6)$$

where  $\hat{\cdot}$  are model predictions,  $\sigma_i$  are metric-wise uncertainties.

The network is relatively small, with around 16,000 parameters, to prevent overfitting. We train the network using an Adam [19] optimiser for 100 epochs with a learning rate of 0.01. Dropout [39] is adopted to regularise the network. We train our model using an 80%–20% train–validation split on Fenland phase 1 data. We test our method on Fenland phase 2 and the smartphone validation study.

### 3.3 Shape Prediction for Smartphone Validation Study

The validation study uses the HKPD [36] method, and generates SMPL avatars using multiview information from RGB images. Given a group of images  $[\mathbf{I}_1, \mathbf{I}_2, \dots]$  of the same participant, each photo is firstly processed using HKPD to generate a Gaussian distribution over SMPL shape parameters  $p(\beta|\mathbf{I}_n)$ . A final body shape

Table 1: Participant characteristics. Data are mean  $\pm$  SD.

		Fenland Study			Smartphone Validation Study
		Phase 1 (training) $n = 9087$	Phase 1 (validation) $n = 2272$	Phase 2 $n = 6102$	$n = 119$
Variables	units				
Age	yrs	48.7 $\pm$ 7.5	48.8 $\pm$ 7.6	55.8 $\pm$ 7.05	42.3 $\pm$ 12.3
Weight	kg	77.6 $\pm$ 15.3	77.6 $\pm$ 15.3	77.3 $\pm$ 15.2	69.9 $\pm$ 13.5
Height	cm	170.0 $\pm$ 9.3	170.0 $\pm$ 9.3	170.4 $\pm$ 9.3	168.4 $\pm$ 9.8
BMI	kg/m <sup>2</sup>	26.8 $\pm$ 4.5	26.7 $\pm$ 4.5	26.6 $\pm$ 4.4	24.6 $\pm$ 4.2
Sex	$n$ (%) male	4298 (47.3)	1035 (45.6)	2979 (48.8)	39 (32.7)
Total body fat mass	kg	26.4 $\pm$ 9.0	26.5 $\pm$ 9.2	26.3 $\pm$ 9.0	21.0 $\pm$ 9.0
Percentage body fat	%	33.5 $\pm$ 7.7	33.8 $\pm$ 8.0	33.7 $\pm$ 7.8	29.4 $\pm$ 9.5
Android fat mass	kg	2.3 $\pm$ 1.2	2.3 $\pm$ 1.2	2.4 $\pm$ 1.2	1.5 $\pm$ 1.1
Gynoid fat mass	kg	4.4 $\pm$ 1.6	4.4 $\pm$ 1.6	4.3 $\pm$ 1.5	3.9 $\pm$ 1.5
Visceral fat mass	kg	1.0 $\pm$ 0.8	1.0 $\pm$ 0.8	1.0 $\pm$ 0.8	0.5 $\pm$ 0.5
Abdominal SCAT mass <sup>a</sup>	kg	1.4 $\pm$ 0.6	1.4 $\pm$ 0.7	1.3 $\pm$ 0.6	1.1 $\pm$ 0.8
Peripheral fat mass <sup>b</sup>	kg	11.6 $\pm$ 4.0	11.7 $\pm$ 4.1	11.3 $\pm$ 4.0	10.2 $\pm$ 4.0
Total lean mass	kg	48.8 $\pm$ 10	48.4 $\pm$ 9.8	48.2 $\pm$ 10	46.8 $\pm$ 10
Appendicular lean mass <sup>c</sup>	kg	22.3 $\pm$ 5.5	22.1 $\pm$ 5.4	21.6 $\pm$ 5.3	21.7 $\pm$ 5.5
ALMI <sup>d</sup>	kg/m <sup>2</sup>	7.6 $\pm$ 1.3	7.5 $\pm$ 1.3	7.4 $\pm$ 1.3	7.5 $\pm$ 1.3

<sup>a</sup> SCAT = subcutaneous adipose tissue

<sup>b</sup> Peripheral fat mass = arms + legs fat mass

<sup>c</sup> Appendicular lean mass = arms + legs lean mass

<sup>d</sup> ALMI: appendicular lean mass index = appendicular lean mass / height<sup>2</sup>

is derived by combining shape information across multiple views according to,

$$p(\beta|\{\mathbf{I}_n\}_{n=1}^N) \propto \prod_{n=1}^N p(\beta|\mathbf{I}_n) \quad (7)$$

where we have assumed conditional independence across views [36, 37].

### 3.4 Statistical Analysis

Statistical analyses were performed using STATA version 17 (StataCorp, College Station, Texas, USA) and Python. A  $P$  value less than 0.05 was considered statistically significant. Descriptive data were reported as mean  $\pm$  standard deviation (SD) or  $n$  (%). Using the methods described in Section 3, we constructed our model to predict total and regional body composition metrics. The performance of the derived model was compared by calculating the Pearson correlation coefficients  $r$  for each outcome parameter and root-mean-square error (RMSE) values. Pearson correlation coefficients were used to investigate associations between the different predicted values of body composition and the measurements of total and regional body composition from DXA. Scatter plots were used to visualise the associations between predicted and measured values. Bland-Altman analysis was used to investigate the agreement between the predicted body composition from our approach against DXA reference measures of total and regional body composition. In the Bland-Altman plot, the y-axis represents the difference or bias between predicted values and measured values (e.g. from





Figure 4: DXA image inputs (Row 1), initial fits using HKPD (Row 2) and optimised fits (Row 3). Initial fits are not accurate enough to represent the participant. Optimised fits have better agreements with the DXA silhouettes.

DXA) with limits of agreement (LoA) described as the 95% confidence range (mean bias  $\pm$  1.96SD), while the x-axis represents the mean value of the reference method (e.g. DXA) rather than the mean between the two methods. Mean differences/biases between the two methods were calculated and significance was tested against zero by paired t-tests. For all the variables, change in body composition was defined as the difference between predictions from Fenland phase 2 (follow up assessment) and Fenland phase 1 (baseline assessment). Scatter plots were used to compare changes in body composition from our predictions with DXA body composition changes. Bland-Altman plots were implemented to assess the agreement between changes in body composition predicted by our method and those measured by DXA. Root-mean-square error (RMSE) was used to assess the accuracy of these comparisons.

## 4 Results

The demographic and anthropometric characteristics of the Fenland study samples and the smartphone validation study are summarised in Table 1. Participants in the smartphone validation study were younger, lighter and leaner, compared to participants in the Fenland study. In terms of body volume, we observed a mean (SD) of 64.4 (13.2) liters in the smartphone validation study, the only dataset with air plethysmography measures.

### 4.1 3D Body Mesh Fitting

Fig. 4 shows samples of our fitted 3D body meshes. Row 1 shows the raw DXA scans. Row 2 shows the initial pose and shape estimations obtained using HKPD [36]. These roughly captures the pose and shape of the body, but the fit to the coronal silhouette is often poor. Row 3 of Fig. 4 shows samples of optimised

fits. We found that optimised meshes agree much better with the silhouettes of DXA images compared with the initial fit. Furthermore, Fig. 8 in the appendix shows samples of meshes before and after optimisation in sagittal view. We observed that the optimisation has resulted in significant changes to the body shapes in the depth dimension when comparing the initial meshes to the optimised meshes. This further shows that initial fits do not represent the actual body shape, and that waist and hip circumferences were needed to generate meshes that better represent the true body shape of participants. Our method, in conclusion, has generated body meshes that are injected with 3D body shape information using paired anthropometry, while creating an improved fit to the DXA silhouettes. SMPL shape parameters corresponding to these optimised meshes are then used for the body composition regressor.

## 4.2 Body Composition Prediction

Table 2 shows the model performance on Fenland and smartphone datasets in the form of mean bias (95% limits of agreement), root-mean-square error (RMSE), and Pearson correlation for the agreement between ground truth body composition values from DXA and model predictions using the 3D meshes. The last column in this table is the model performance using the smartphone method. Fig. 9 in the appendix shows selected scatter plots of model predictions against target values on Fenland phase 2 data. Fig. 5 shows scatter plots and Bland-Altman plots for percentage body fat for Fenland phase 1 (validation), phase 2 (validation), smartphone study using DXA silhouette optimisation, and smartphone study using RGB images.

In the Fenland phase 1 validation sample, correlation coefficients between predicted and measured DXA parameters were strong ( $r > 0.89$ ) for all fat mass and lean mass variables. Bland-Altman analyses revealed no significant mean bias for the following predicted DXA parameters: fat mass, android fat, appendicular lean mass, percentage body fat and appendicular lean mass index (all  $P > 0.05$ ). However, significant ( $P < 0.05$ ) mean bias was observed for total lean mass, visceral fat, abdominal SCAT mass, gynoid fat, and peripheral fat: mean biases (95% LoA) were 0.11 (−4.7; 4.9)kg, −0.03 (−0.7; 0.6)kg, 0.05 (−0.5; 0.6)kg, 0.04 (−1; 1)kg, −0.06 (−2.7; 2.6)kg respectively. In the Fenland phase 2 validation sample, which included older individuals, correlation coefficients between predicted and measured DXA parameters were also strong ( $r > 0.86$ ) for all the body composition variables. Agreement analyses revealed significant ( $P < 0.05$ ) mean bias for all predicted DXA measured parameters. Mean bias was −0.24 (−6.7; 6.2)% for percentage body fat and for the other body composition metrics, the mean bias ranged between −0.04 to 0.41 kg. Similar results for the DXA silhouette method were also observed in the external validation (Column 3 of Table 2), which included younger individuals. Correlation coefficients between predicted and measured DXA parameters were  $r > 0.89$ . Mean bias was 1.13 (−7.2; 9.5)% for percentage body fat and for the other body composition parameters, the mean bias ranged between −1.21 to 1.13kg.

Results using SMPL shape generated directly from the four smartphone images in this external validation is shown in Column 4 of Table 2. The correlations coefficients between DEXA metrics and all the predicted body composition values were  $r > 0.84$ . Mean bias (95% LoA) was 1.62 (−9.2; 12.5)% for percentage body fat and for the other body composition parameters, the mean bias ranged between −1.51 to 0.93kg. Volume derived from the smartphone achieved RMSE of 5.21liters compared to BODPOD volume, with mean bias (95% LoA) of 5.36 (−4.9; 15.6)liters. We also compared accuracy of waist, hip, calf, and arm circumferences from the smartphone avatars. Mean biases were 2.49 (−7.4; 12.4)cm, 2.25 (−6.3; 10.8)cm, −4.09 (1.3; 6.7)cm, and −3.03 (−7.8; 1.7)cm respectively. RMSE were 5.04cm, 4.38cm, 2.76cm and 2.42cm respectively.

Table 2: Agreement analysis of 3D body shape predicted vs measured body composition.

	Fenland phase 1 (validation)			Fenland phase 2 (validation)			Smartphone Validation Study						
	DXA silhouette $n = 2272$			DXA silhouette $n = 6102$			DXA silhouette $n = 119$			4 phone photos $n = 119$			
	Bias (95%LoA)	RMSE	$r^a$	Bias (95%LoA)	RMSE	$r^a$	Bias (95%LoA)	RMSE	$r^a$	Bias (95%LoA)	RMSE	$r^a$	
DXA													
	Total body fat mass (kg)	-0.09 (-4.9; 4.7)	2.45	0.96	-0.13 (-5; 4.8)	2.50	0.96	0.57 (-5; 6.2)	2.86	0.95	0.93 (-6.4; 8.3)	3.76	0.91
	Percentage body fat (%)	-0.07 (-6.4; 6.2)	3.22	0.91	-0.24 (-6.7; 6.2)	3.28	0.91	1.13 (-7.2; 9.5)	4.26	0.90	1.62 (-9.2; 12.5)	5.54	0.84
	Android fat mass (kg)	0 (-0.7; 0.8)	0.38	0.95	0.03 (-0.7; 0.8)	0.38	0.95	0.07 (-0.6; 0.8)	0.36	0.95	0.12 (-0.9; 1.1)	0.50	0.89
	Gynoid fat mass (kg)	0.04 (-1; 1)	0.51	0.95	0.02 (-1; 1)	0.51	0.94	0.06 (-1.1; 1.2)	0.59	0.92	0.13 (-1.2; 1.3)	0.67	0.90
	Visceral fat mass (kg)	-0.03 (-0.7; 0.6)	0.34	0.89	-0.04 (-0.8; 0.7)	0.37	0.90	-0.02 (-0.4; 0.4)	0.19	0.94	0.08 (-0.5; 0.7)	0.29	0.85
	Abdominal SCAT mass (kg) <sup>b</sup>	0.05 (-0.5; 0.6)	0.30	0.90	0.09 (-0.5; 0.7)	0.31	0.86	0.07 (-0.7; 0.9)	0.42	0.89	0.03 (-0.9; 1.0)	0.49	0.86
	Peripheral fat mass (kg) <sup>c</sup>	-0.06 (-2.7; 2.6)	1.36	0.94	-0.11 (-2.8; 2.6)	1.39	0.93	0.12 (-2.6; 2.9)	1.40	0.94	0.44 (-3.0; 3.9)	1.74	0.91
	Total lean mass (kg)	0.11 (-4.7; 4.9)	2.45	0.97	0.32 (-4.7; 5.3)	2.54	0.97	-0.95 (-6.9; 5)	3.05	0.95	-1.51 (-9.0; 6.0)	3.82	0.93
	Appendicular lean mass (kg) <sup>d</sup>	-0.01 (-2.8; 2.7)	1.40	0.97	0.41 (-2.3; 3.2)	1.40	0.97	-1.21 (-4.2; 1.8)	1.53	0.96	-1.25 (-5.3; 2.8)	2.05	0.93
	ALMI (kg/m <sup>2</sup> ) <sup>e</sup>	-0.01 (-0.9; 0.9)	0.48	0.93	0.14 (-0.8; 1.1)	0.48	0.93	-0.43 (-1.5; 0.6)	0.54	0.91	-0.42 (-1.8; 1.0)	0.70	0.84
BODPOD	Volume (L)	-	-	-	-	-	-	-	-	-	5.36 (-4.9; 15.6)	5.21	0.91
Anthropometry	Waist Circumference (cm)	-	-	-	-	-	-	-	-	-	2.49 (-7.4; 12.4)	5.04	0.89
	Hip Circumference (cm)	-	-	-	-	-	-	-	-	-	2.25 (-6.3; 10.8)	4.38	0.87
	Calf circumference (cm)	-	-	-	-	-	-	-	-	-	-4.09 (1.3; 6.7)	2.76	0.59
	Arm Circumference (cm)	-	-	-	-	-	-	-	-	-	-3.03 (-7.8; 1.7)	2.42	0.80

<sup>a</sup>  $r$  = Pearson's correlations

<sup>b</sup> SCAT = subcutaneous adipose tissue

<sup>c</sup> Peripheral fat mass = arms + legs fat mass

<sup>d</sup> Appendicular lean mass = arms + legs lean mass

<sup>e</sup> ALMI: appendicular lean mass index = appendicular lean mass / height<sup>2</sup>

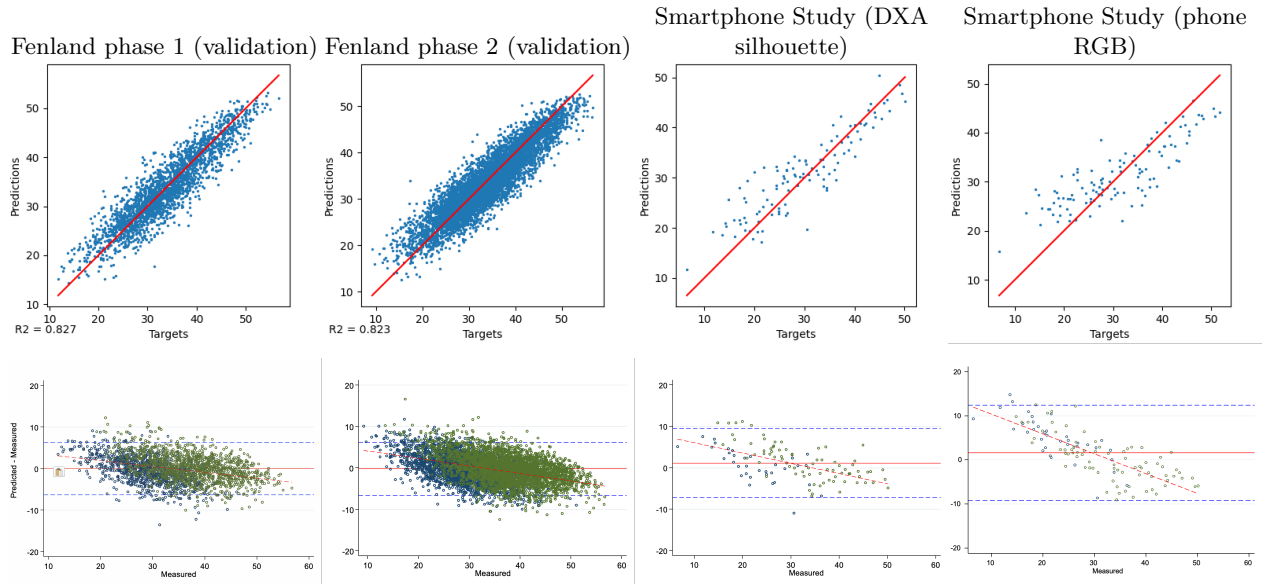


Figure 5: Scatter plots (Row 1) and Bland-Altman plots (Row 2) for agreement between predicted and measured percentage body fat in the Fenland samples and the external validation set. All predictions are based on 3D meshes derived from DXA silhouettes, except last panel which is based on 3D meshes derived from four RGB photos from smartphone.

### 4.3 Comparison between Prediction Models

We conducted a comparison study on different regressor model inputs, to verify that 3D body meshes provide crucial information for the downstream composition regressor. Results of the different models investigated in the comparison study are shown in Table 3. Model A (weight and height only) achieved some level of predictive ability. Performance of the model was improved by adding waist and hip circumferences (Models B and C), as waist and hip are strong indicators of composition metrics such as android fat mass and gynoid fat mass. In the final model (Model E), we quantified the contribution of the SMPL shape parameters in addition to using height and weight only. This model substantially improved the estimation of the body composition metrics compared to anthropometry alone. The explained variance ( $R^2$ ) in percentage body fat increased from 73% to 82%; total fat mass from 88% to 92%; total lean mass from 91% to 93%; gynoid fat from 81% to 89%; android fat from 81% to 89%; peripheral fat mass from 80% to 87%; appendicular lean mass 90% to 93%; appendicular lean mass index from 74% to 86%; visceral fat from 70% to 80% and abdominal SCAT from 70% to 72%. We also attempted to predict body composition using a simple linear regressor (Model D), but the neural network approach (Model E) outperforms it noticeably.

Table 3: Comparison study on model inputs. Best R-squared values are in bold font. We verify that the addition of shape parameters is crucial for the model to perform well. Our model using SMPL outperforms the other models in almost every metric.

Model Name	A	B	C	D	E
Model Inputs	-	-	-	SMPL	SMPL
	Height, Weight	Height, Weight	Height, Weight	Height, Weight	Height, Weight
	-	Waist	Waist, Hip	-	-
Method	Network	Network	Network	Linear	Network
Metrics	Fenland phase 2 $R^2$				
Total fat mass	0.884	0.909	0.910	0.897	<b>0.922</b>
Percentage body fat	0.739	0.792	0.797	0.766	<b>0.823</b>
Android fat mass	0.809	0.887	0.884	0.880	<b>0.894</b>
Gynoid fat mass	0.811	0.830	0.863	0.843	<b>0.886</b>
Visceral fat mass	0.698	0.779	0.792	0.774	<b>0.802</b>
Abdominal SCAT mass <sup>a</sup>	0.701	0.727	<b>0.730</b>	0.716	0.723
Peripheral fat mass <sup>b</sup>	0.802	0.824	0.832	0.821	<b>0.872</b>
Total lean mass	0.910	0.925	0.921	0.916	<b>0.934</b>
Appendicular lean mass <sup>c</sup>	0.895	0.921	0.911	0.906	<b>0.927</b>
ALMI <sup>d</sup>	0.739	0.853	0.824	0.833	<b>0.853</b>

<sup>a</sup> SCAT = subcutaneous adipose tissue

<sup>b</sup> Peripheral fat mass = arms + legs fat mass

<sup>c</sup> Appendicular lean mass = arms + legs lean mass

<sup>d</sup> ALMI: appendicular lean mass index = appendicular lean mass / height<sup>2</sup>

#### 4.4 Predictions of Body Composition Changes

A total of 5733 individuals participated in both Fenland phase 1 and phase 2, which enabled us to examine the model’s ability to detect within-individual body composition changes over a mean (SD) of 6.7 (2.0) years.

Table 4 shows model predictions of changes in body composition metrics for individuals present in both Fenland phases. Our model was able to detect change for numerous fat mass metrics. The agreement between predicted body composition values and DXA parameters ( $r$ ) for changes in percentage body fat, total fat mass, gynoid fat mass, android fat mass, peripheral fat mass, visceral fat mass and abdominal SCAT mass were 0.92, 0.76, 0.87, 0.83, 0.74, 0.76, 0.82 respectively, with RMSE of 2.3%, 1.75kg, 0.39kg, 0.29kg, 1.00kg, 0.26kg, 0.21kg respectively. Changes in lean mass were less well captured, mainly due to the fact that lean mass largely remains unchanged for most individuals over this time period.  $r$  values for change for total lean mass, appendicular lean mass and ALMI were 0.60, 0.64, 0.63 respectively, with RMSE of 1.82kg, 1.06kg, 0.36kg respectively. Fig. 6 shows selected scatter plots between the predicted changes in percentage body fat, lean mass, android and gynoid fat mass against changes measured by DXA for the same variables and the corresponding Bland-Altman plots.

## 5 Discussion

In this paper, we derived a novel computer vision-based method that fits a 3D body mesh to a single DXA silhouette with paired anthropometry data (height, waist and hip circumferences). Using our method, we generated a large database of 3D body meshes ( $n = 17,461$ ) with paired anthropometric and metabolic traits. We then showed that total and regional body composition metrics could be predicted accurately using these meshes. In the comparison study, we showed that shape parameters provide additional cues for predicting

Table 4: Predictions of composition changes for participants in both Fenland phases. Our model is able to predict meaningful changes in body fat masses.

Metrics	Bias (95%LoA)	RMSE	$r$
Total fat mass	-0.11 (-3.5; 3.3)	1.75	0.92
Percentage body fat	-0.25 (-4.7; 4.2)	2.30	0.76
Android fat mass	0.01 (-0.6; 0.6)	0.29	0.87
Gynoid fat mass	-0.02 (-0.8; 0.7)	0.39	0.83
Visceral fat mass	-0.02 (-0.5; 0.5)	0.26	0.74
Abdominal SCAT mass <sup>a</sup>	0.03 (-0.4; 0.4)	0.21	0.76
Peripheral fat mass <sup>b</sup>	-0.05 (-2.0; 1.9)	1.00	0.82
Total lean mass	0.32 (-3.2; 3.8)	1.82	0.60
Appendicular lean mass <sup>c</sup>	0.45 (-1.4; 2.3)	1.06	0.64
ALMI <sup>d</sup>	0.16 (-0.5; 0.8)	0.36	0.63

<sup>a</sup> SCAT = subcutaneous adipose tissue

<sup>b</sup> Peripheral fat mass = arms + legs fat mass

<sup>c</sup> Appendicular lean mass = arms + legs lean mass

<sup>d</sup> ALMI: appendicular lean mass index = appendicular lean mass / height<sup>2</sup>

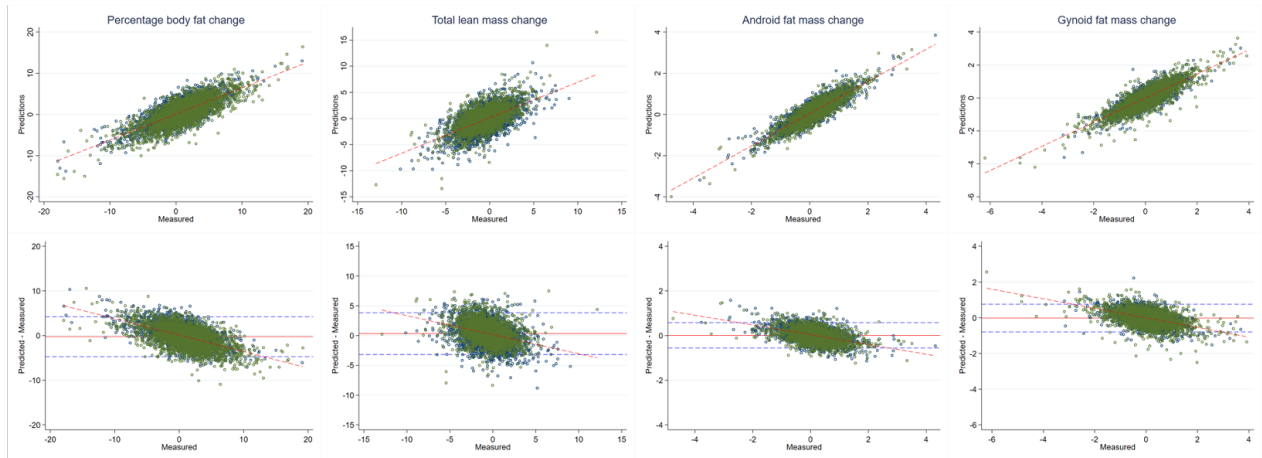


Figure 6: Selected scatter plots (Row 1) and corresponding Bland-Altman plots (Row 2) for prediction of change in body composition metrics.

body composition by comparing the performance of models with different inputs. We demonstrated the derived model’s ability to detect longitudinal change in these characteristics over time by making predictions for individuals that were present in both phases of the Fenland study. In our smartphone validation study, we showed how avatars were generated using four smartphone photographs directly, without the use of optimisation. Finally, we showed preliminary body composition prediction results using these avatars.

The model using optimised 3D meshes predicted body composition metrics with sufficient accuracy to assess relative differences between individuals, and was sufficiently accurate to predict absolute values for total and regional body composition including visceral fat and abdominal SCAT mass, as well as lean mass, appendicular lean mass, and ALMI. All the body composition metrics predicted from the optimised 3D meshes showed Pearson correlation coefficients  $r > 0.86$ . In a similar study, Xie et al. [48] which used a 2D PCA shape space constructed from keypoints selected on a silhouette achieved the following results on a cohort of size 1,554: percentage body fat  $R^2$  adj. (adjusted  $R^2$ ) = 0.728, RMSE = 3.12% for boys, and  $R^2$  adj. = 0.691, RMSE = 3.39% for girls. However, this comparison is limited as this study was conducted in children, as they have proportionally larger body surface area to volume ratio than adults, as well as sex differences in dimensions of body shapes such as central:peripheral ratio [34]. In a more relevant study, Ng

et al. [29] which predicted composition using PCA parameters of 3DO scans on a cohort of size 407, observed similar results to our analysis: total fat mass  $R^2 = 0.91$ , RMSE = 3.07kg for males,  $R^2 = 0.95$ , RMSE = 2.63kg for females; percentage body fat  $R^2 = 0.70$ , RMSE = 3.55% for males,  $R^2 = 0.72$ , RMSE = 3.88% for females. Similarly, Tian et al. [41] fitted a 3D shape to a single coronal silhouette on a cohort of size 416, and predicted body composition with the following results: total fat mass  $R^2 = 0.90$ , RMSE = 3.63kg for males,  $R^2 = 0.94$ , RMSE = 2.29kg for females; percentage body fat:  $R^2 = 0.725$ , RMSE = 3.90% for male,  $R^2 = 0.74$ , RMSE = 3.29% for female. In comparison, we achieved similar performance on a much larger dataset  $n = 6102$ , with the following results: total fat mass  $R^2 = 0.922$ , RMSE = 2.5kg; percentage body fat  $R^2 = 0.823$ , RMSE = 3.28% overall. In addition, our approach also injects information in the depth dimension by using waist and hip circumferences, while the fitting by Tian et al. [41] is limited to the coronal silhouette. In Klarqvist et al. [23], stronger correlations were observed for visceral fat, abdominal SCAT mass and gynoid fat, using coronal and sagittal silhouettes derived from MRI, the gold standard for those measures of adiposity. Our estimates for these metrics were based on an in-built algorithm from the DXA manufacturer, which is not a criterion method. However, compared to Klarqvist et al. [23], our analysis included more body composition metrics such as appendicular lean mass and its index, which are used as a proxy for the assessment of sarcopenia [22, 46]. Even though comparison to these studies may be limited as they were conducted in cohorts of different ages, using different body composition instruments and computer vision approaches, we have shown that our method produced meshes that were comparably accurate to 3DO scans [29], by predicting body composition to a similar accuracy on a large test dataset. This allows for the construction of large 3D shape databases using our method, and in turn enables larger deep learning models to be used to analyse 3D body shapes. Our prediction results for change in body composition were similar to Wong et al. [47] which report predicting DXA metrics from 3DO scanning images from 133 participants, where the change in fat mass is slightly underestimated, and lean mass overestimated.

The strengths of our study include the large sample size of the Fenland study, the same DXA instruments in the different samples, the same DXA analytical software, and the robust validation in two separate independent cohorts (Fenland phase 2, representing an older group from the derivation sample and the external validation, which consisted of younger individuals). Furthermore, our method assesses changes to body composition over time. Another benefit of the 3D body mesh approach is to enable anonymity of user data. Our method does not require raw images of the participants to be retained, rather we only store the generated body meshes, which could also be done efficiently by only storing the SMPL shape parameters. This provides additional incentive in scientific and clinical studies for participants to partake in data collection, since concern over sharing sensitive information is largely eliminated. Compared to other smartphone apps that estimate body composition from photographs, we note that they either do not reconstruct 3D meshes [27, 40], or they require strict pose constraints [28, 40]. Our app works robustly for noisy backgrounds, can be extended to incorporate an arbitrary number of images, and in practice works for arbitrary body and camera poses.

This work is not without limitations. We acknowledge that our study samples were predominantly adults of white European origin. Future analyses should assess the validity of these models in other ethnic groups as well as younger populations, since there are significant racial and age differences in body composition [9, 13, 16, 30, 34]. Future work should also focus on improving the avatar accuracy generated using the smartphone app. While our findings support the validity of our method in Fenland phase 2 data and in the smartphone validation study, we expected and found lower performance using avatars derived from RGB images, as avatars obtained using the phone app can be inaccurate. We do not optimise our smartphone-generated avatars although this would improve prediction accuracy, since we do not retain the photographs due to ethical constraints and data security. Alternatively, using domain-agnostic representations such as waist-hip ratio, waist-height ratio might produce stronger results as they are normalised with respect to height. Choudhary et al. [8] showed that accurate waist-hip ratio could be derived directly from images using attention based networks, which could prove useful in this regard. With improvements to avatar accuracy, app-generated avatars would be able to approach the prediction performance on Fenland.

Through the implementation of the app, users will be able to visualize their body shapes, and track potential changes using a portable and relatively inexpensive but accurate device. Using our method in clinical research studies, we could potentially identify individuals at the highest risk of preventable complications (e.g. significant increase in body fat from their first assessment). For instance, Fig. 7 shows the modeled body

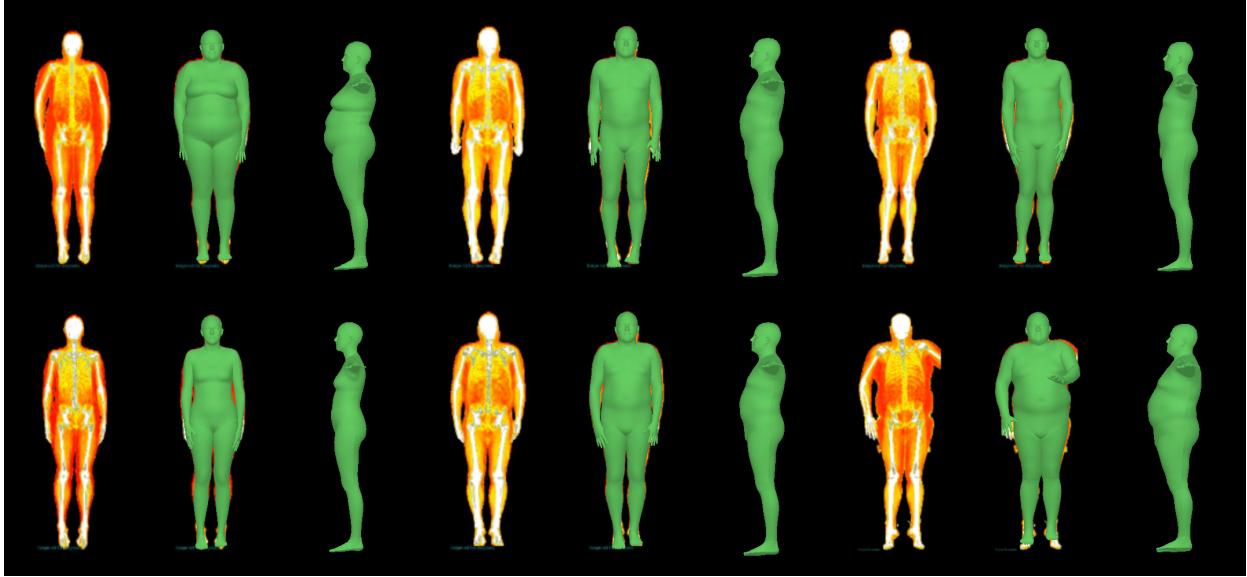


Figure 7: Each three columns show one participant who lost (Column 1–3), maintained (Column 4–6) or gained (Column 7–9) fat mass, their DXA scans and fitted meshes for Fenland phase 1 (Row 1) and phase 2 (Row 2). Changes in body shape for the first and third participant is significant.

shapes at phase 1 and 6 years later at phase 2 of three participants who either maintained, gained, or lost fat mass. The first and last three columns show two participants who lost and gained lots of fat mass between the two phases. Significant changes could be seen comparing the two sets of meshes. The end goal would be to encourage users to adopt a healthier way of life, by visualizing changes to their body shapes with time, rather than just focusing on numerical values (e.g. increase in BMI).

In conclusion, capturing two-dimensional (2D) images coupled with appropriate inference techniques to reconstruct a 3D model of the body, may be a viable alternative tool to clinical medical imaging, and it could offer a readily accessible health metric for monitoring the efficacy of lifestyle interventions.

## Funding Statement and Acknowledgements

We thank all study volunteers who participated in the Fenland and validation studies as well as the research teams at MRC Epidemiology Unit and NIHR Cambridge Clinical Research Facility responsible for data collection and management. We also acknowledge support from the Medical Research Council (Unit programmes MC\_UU\_12015/1, MC\_UU\_12015/2, MC\_UU\_12015/3, MC\_UU\_12015/4 and MC\_UU\_12015/5). The Fenland Study was funded by the Medical Research Council and the Wellcome Trust. We are grateful to funding to Powell and De Lucia Rolfe, who are supported by the NIHR Biomedical Research Centre Cambridge [IS-BRC-1215-20014]. The NIHR Cambridge Biomedical Research Centre (BRC) is a partnership between Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge, funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. Watson is supported by the NIHR Cambridge Clinical Research Facility.

## Author Contributions

Qiao had full access to all of the data used in this analysis and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Qiao, De Lucia Rolfe, Mak, Brage, Cipolla; Acquisition, analysis, or interpretation of data: Qiao, De Lucia Rolfe, Mak, Sengupta, Powell, Watson, Heymsfield, Shepherd, Wareham, Brage, Cipolla; Drafting of the manuscript: Qiao, De Lucia Rolfe, Brage, Cipolla; Critical revision of the manuscript for important intellectual content: Qiao, De Lucia Rolfe, Mak, Sengupta, Powell, Watson, Heymsfield, Shep-



herd, Wareham, Brage, Cipolla; Statistical analysis: Qiao, De Lucia Rolfe, Powell; Obtained funding: Brage, Wareham, Watson; Administrative, technical, or material support: Qiao, De Lucia Rolfe, Powell, Watson; Supervision: Brage, Cipolla.

## References

- [1] Diet, anthropometry and physical activity (dapa) measurement toolkit. [www.measurement-toolkit.org](http://www.measurement-toolkit.org). Accessed: 2024-04-02. 4
- [2] Richard N Baumgartner, Steven B Heymsfield, and Alex F Roche. Human body composition and the epidemiology of chronic disease. *Obesity research*, 3(1):73–95, 1995. 1
- [3] Jonathan P Bennett, Yong En Liu, Brandon K Quon, Nisa N Kelly, Lambert T Leong, Michael C Wong, Samantha F Kennedy, Dominic C Chow, Andrea K Garber, Ethan J Weiss, et al. Three-dimensional optical body shape and features improve prediction of metabolic disease risk in a diverse sample of adults. *Obesity*, 30(8):1589–1598, 2022. 2
- [4] Jonathan P Bennett, Yong En Liu, Brandon K Quon, Nisa N Kelly, Michael C Wong, Samantha F Kennedy, Dominic C Chow, Andrea K Garber, Ethan J Weiss, Steven B Heymsfield, et al. Assessment of clinical measures of total and regional body composition from a commercial 3-dimensional optical body scanner. *Clinical Nutrition*, 41(1):211–218, 2022. 2
- [5] Federica Bogo, Angjoo Kanazawa, Christoph Lassner, Peter Gehler, Javier Romero, and Michael J Black. Keep it smpl: Automatic estimation of 3d human pose and shape from a single image. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part V 14*, pages 561–578. Springer, 2016. 5
- [6] Magnus Borga, Janne West, Jimmy D Bell, Nicholas C Harvey, Thobias Romu, Steven B Heymsfield, and Olof Dahlqvist Leinhard. Advanced body composition assessment: from body mass index to body composition profiling. *Journal of Investigative Medicine*, 66(5):1–9, 2018. 1
- [7] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O’Connell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, 2018. 2
- [8] Siddharth Choudhary, Ganesh Iyer, Brandon M Smith, Jinjin Li, Mark Sippel, Antonio Criminisi, and Steven B Heymsfield. Development and validation of an accurate smartphone application for measuring waist-to-hip circumference ratio. *NPJ Digital Medicine*, 6(1):168, 2023. 15
- [9] Joan M Conway, Susan Z Yanovski, Nib A Avila, and Van S Hubbard. Visceral adipose tissue differences in black and white women. *The American journal of clinical nutrition*, 61(4):765–771, 1995. 15
- [10] Joshua J Fenton and Richard A Deyo. Patient self-referral for radiologic screening tests: clinical and ethical concerns. *The Journal of the American Board of Family Practice*, 16(6):494–501, 2003. 1
- [11] Austin J Graybeal, Caleb F Brandner, and Grant M Tinsley. Evaluation of automated anthropometrics produced by smartphone-based machine learning: a comparison with traditional anthropometric assessments. *British Journal of Nutrition*, 130(6):1077–1087, 2023. 2
- [12] J Haarbo, A Gotfredsen, C Hassager, and CJCP Christiansen. Validation of body composition by dual energy x-ray absorptiometry (dexa). *Clinical Physiology*, 11(4):331–341, 1991. 1
- [13] James O Hill, Steve Sidney, Cora E Lewis, Kim Tolan, Ann L Scherzinger, and Elizabeth R Stamm. Racial differences in amounts of visceral adipose tissue in young adults: the cardia (coronary artery risk development in young adults) study. *The American journal of clinical nutrition*, 69(3):381–387, 1999. 15
- [14] Judy Illes, John E Desmond, Lynn F Huang, Thomas A Raffin, and Scott W Atlas. Ethical and practical considerations in managing incidental findings in functional magnetic resonance imaging. *Brain and cognition*, 50(3):358–365, 2002. 1

- [15] Henry S Kahn, Qiuping Gu, Kai McKeever Bullard, David S Freedman, Namanjeet Ahluwalia, and Cynthia L Ogden. Population distribution of the sagittal abdominal diameter (sad) from a representative sample of us adults: comparison of sad, waist circumference and body mass index for identifying dysglycemia. PloS one, 9(10):e108707, 2014. [6](#)
- [16] JA Kanaley, Ifigenia Giannopoulou, G Tillapaugh-Fay, JS Nappi, and LL Ploutz-Snyder. Racial differences in subcutaneous and visceral fat distribution in postmenopausal black and white women. Metabolism, 52(2):186–191, 2003. [15](#)
- [17] Angjoo Kanazawa, Michael J Black, David W Jacobs, and Jitendra Malik. End-to-end recovery of human shape and pose. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7122–7131, 2018. [5](#)
- [18] Alex Kendall, Yarin Gal, and Roberto Cipolla. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7482–7491, 2018. [7](#)
- [19] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014. [7](#)
- [20] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114, 2013. [2](#)
- [21] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 4015–4026, 2023. [2](#)
- [22] Akihiko Kitamura, Satoshi Seino, Takumi Abe, Yu Nofuji, Yuri Yokoyama, Hidenori Amano, Mariko Nishi, Yu Taniguchi, Miki Narita, Yoshinori Fujiwara, et al. Sarcopenia: prevalence, associated factors, and the risk of mortality and disability in japanese older adults. Journal of cachexia, sarcopenia and muscle, 12(1):30–38, 2021. [4](#), [15](#)
- [23] Marcus DR Klarqvist, Saaket Agrawal, Nathaniel Diamant, Patrick T Ellinor, Anthony Philippakis, Kenney Ng, Puneet Batra, and Amit V Khera. Silhouette images enable estimation of body fat distribution and associated cardiometabolic risk. NPJ digital medicine, 5(1):105, 2022. [2](#), [15](#)
- [24] Lambert T Leong, Michael C Wong, Yong E Liu, Yannik Glaser, Brandon K Quon, Nisa N Kelly, Devon Cataldi, Peter Sadowski, Steven B Heymsfield, and John A Shepherd. Generative deep learning furthers the understanding of local distributions of fat and muscle on body shape and health using 3d surface scans. Communications Medicine, 4(1):13, 2024. [2](#)
- [25] T Lindsay, K Westgate, K Wijndaele, S Hollidge, N Kerrison, N Forouhi, et al. Descriptive epidemiology of physical activity energy expenditure in uk adults. The Fenland Study. medRxiv, 1:19003442, 2019. [3](#)
- [26] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. In Seminal Graphics Papers: Pushing the Boundaries, Volume 2, pages 851–866. 2023. [5](#)
- [27] Maulik D Majmudar, Siddhartha Chandra, Kiran Yakkala, Samantha Kennedy, Amit Agrawal, Mark Sippel, Prakash Ramu, Apoorv Chaudhri, Brooke Smith, Antonio Criminisi, et al. Smartphone camera based assessment of adiposity: a validation study. NPJ Digital Medicine, 5(1):79, 2022. [2](#), [15](#)
- [28] Cassidy McCarthy, Grant M Tinsley, Shengping Yang, Brian A Irving, Michael C Wong, Jonathan P Bennett, John A Shepherd, and Steven B Heymsfield. Smartphone prediction of skeletal muscle mass: model development and validation in adults. The American Journal of Clinical Nutrition, 117(4):794–801, 2023. [2](#), [15](#)

- [29] Bennett K Ng, Markus J Sommer, Michael C Wong, Ian Pagano, Yilin Nie, Bo Fan, Samantha Kennedy, Brianna Bourgeois, Nisa Kelly, Yong E Liu, et al. Detailed 3-dimensional body shape features predict body composition, blood metabolites, and functional strength: the shape up! studies. The American journal of clinical nutrition, 110(6):1316–1326, 2019. 2, 15
- [30] Arlette C Perry, E Brooks Applegate, M Loreto Jackson, Steven Deprima, Ronald B Goldberg, Robert Ross, Lani Kempner, and Brandon B Feldman. Racial differences in visceral adipose tissue but not anthropometric markers of health-related variables. Journal of applied physiology, 89(2):636–643, 2000. 15
- [31] Andrew M Prentice and Susan A Jebb. Beyond body mass index. Obesity reviews, 2(3):141–147, 2001. 2
- [32] Nikhila Ravi, Jeremy Reizenstein, David Novotny, Taylor Gordon, Wan-Yen Lo, Justin Johnson, and Georgia Gkioxari. Accelerating 3d deep learning with pytorch3d. arXiv preprint arXiv:2007.08501, 2020. 6
- [33] Cornelius Johannes Frederik Reyneke, Marcel Lüthi, Valérie Burdin, Tania S Douglas, Thomas Vetter, and Tinashe EM Mutsvangwa. Review of 2-d/3-d reconstruction using statistical shape and intensity models and x-ray image synthesis: Toward a unified framework. IEEE reviews in biomedical engineering, 12:269–286, 2018. 5
- [34] Leonardo P Santos, Ken K Ong, Felix Day, Jonathan CK Wells, Alicia Matijasevich, Ina S Santos, Cesar G Victora, and Aluisio JD Barros. Body shape and size in 6-year old children: assessment by three-dimensional photonic scanning. International Journal of Obesity, 40(6):1012–1017, 2016. 14, 15
- [35] Akash Sengupta, Ignas Budvytis, and Roberto Cipolla. Synthetic training for accurate 3d human pose and shape estimation in the wild. In British Machine Vision Conference (BMVC), September 2020. 5
- [36] Akash Sengupta, Ignas Budvytis, and Roberto Cipolla. Hierarchical kinematic probability distributions for 3d human shape and pose estimation from images in the wild. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 11219–11229, 2021. 2, 5, 7, 8, 9
- [37] Akash Sengupta, Ignas Budvytis, and Roberto Cipolla. Probabilistic 3d human shape and pose estimation from multiple unconstrained images in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 16094–16104, 2021. 2, 5, 8
- [38] Brooke Smith, Cassidy McCarthy, Marcelline E Dechenaud, Michael C Wong, John Shepherd, and Steven B Heymsfield. Anthropometric evaluation of a 3d scanning mobile application. Obesity, 30(6):1181–1188, 2022. 2
- [39] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1):1929–1958, 2014. 7
- [40] Katherine Sullivan, Casey J Metoyer, Bjoern Hornikel, Clifton J Holmes, Brett S Nickerson, Michael R Esco, and Michael V Fedewa. Agreement between a 2-dimensional digital image-based 3-compartment body composition model and dual energy x-ray absorptiometry for the estimation of relative adiposity. Journal of Clinical Densitometry, 25(2):244–251, 2022. 2, 15
- [41] Isaac Y Tian, Bennett K Ng, Michael C Wong, Samantha Kennedy, Phoenix Hwaung, Nisa Kelly, En Liu, Andrea K Garber, Brian Curless, Steven B Heymsfield, et al. Predicting 3d body shape and body composition from conventional 2d photography. Medical Physics, 47(12):6232–6245, 2020. 2, 15
- [42] Isaac Y Tian, Michael C Wong, Samantha Kennedy, Nisa N Kelly, Yong E Liu, Andrea K Garber, Steven B Heymsfield, Brian Curless, and John A Shepherd. A device-agnostic shape model for automated body composition estimates from 3d optical scans. Medical physics, 49(10):6395–6409, 2022. 2

- [43] Grant M Tinsley, M Lane Moore, Marqui L Benavides, Jacob R Dellinger, and Brian T Adamson. 3-dimensional optical scanning for body composition assessment: A 4-component model comparison of four commercially available scanners. Clinical Nutrition, 39(10):3160–3167, 2020. [2](#)
- [44] Antti Tolonen, Tomppa Pakarinen, Antti Sassi, Jere Kyttä, William Cancino, Irina Rinta-Kiikka, Said Pertuz, and Otso Arponen. Methodology, clinical applications, and future directions of body composition analysis using computed tomography (ct) images: a review. European journal of radiology, 145:109943, 2021. [1](#)
- [45] David R Weber, René H Moore, Mary B Leonard, and Babette S Zemel. Fat and lean bmi reference curves in children and adolescents and their utility in identifying excess adiposity compared with bmi and percentage body fat. The American journal of clinical nutrition, 98(1):49–56, 2013. [2](#)
- [46] Zhangxin Wen, Jiaxuan Gu, Rong Chen, Qinyi Wang, Na Ding, Lingqiong Meng, Xiangbing Wang, Hong Liu, Zhifeng Sheng, and Houfeng Zheng. Handgrip strength and muscle quality: results from the national health and nutrition examination survey database. Journal of Clinical Medicine, 12(9):3184, 2023. [15](#)
- [47] Michael C Wong, Jonathan P Bennett, Lambert T Leong, Isaac Y Tian, Yong E Liu, Nisa N Kelly, Cassidy McCarthy, Julia MW Wong, Cara B Ebbeling, David S Ludwig, et al. Monitoring body composition change for intervention studies with advancing 3d optical imaging technology in comparison to dual-energy x-ray absorptiometry. The American Journal of Clinical Nutrition, 117(4):802–813, 2023. [15](#)
- [48] Bowen Xie, Jesus I Avila, Bennett K Ng, Bo Fan, Victoria Loo, Vicente Gilsanz, Thomas Hangartner, Heidi J Kalkwarf, Joan Lappe, Sharon Oberfield, et al. Accurate body composition measures from whole-body silhouettes. Medical physics, 42(8):4668–4677, 2015. [2](#), [14](#)
- [49] Weimin Yu, Moritz Tannast, and Guoyan Zheng. Non-rigid free-form 2d–3d registration using a b-spline-based statistical deformation model. Pattern recognition, 63:689–699, 2017. [5](#)

## 6 Appendix

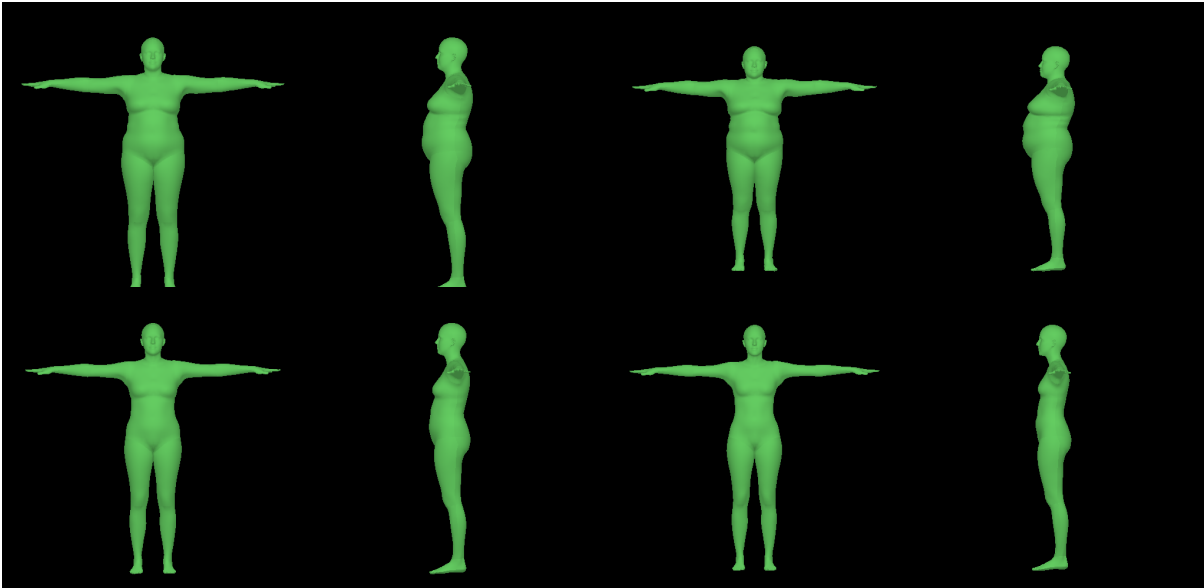


Figure 8: Column 1, 2: Initial body shape fits using HKPD (in T-pose, front and side view); Column 3, 4: Optimised fits (in T-pose, front and side view). Initial fits lack information in the depth axis. Using anthropometry (height, waist, hip circumferences), optimised body meshes reflect the actual body shape much better.

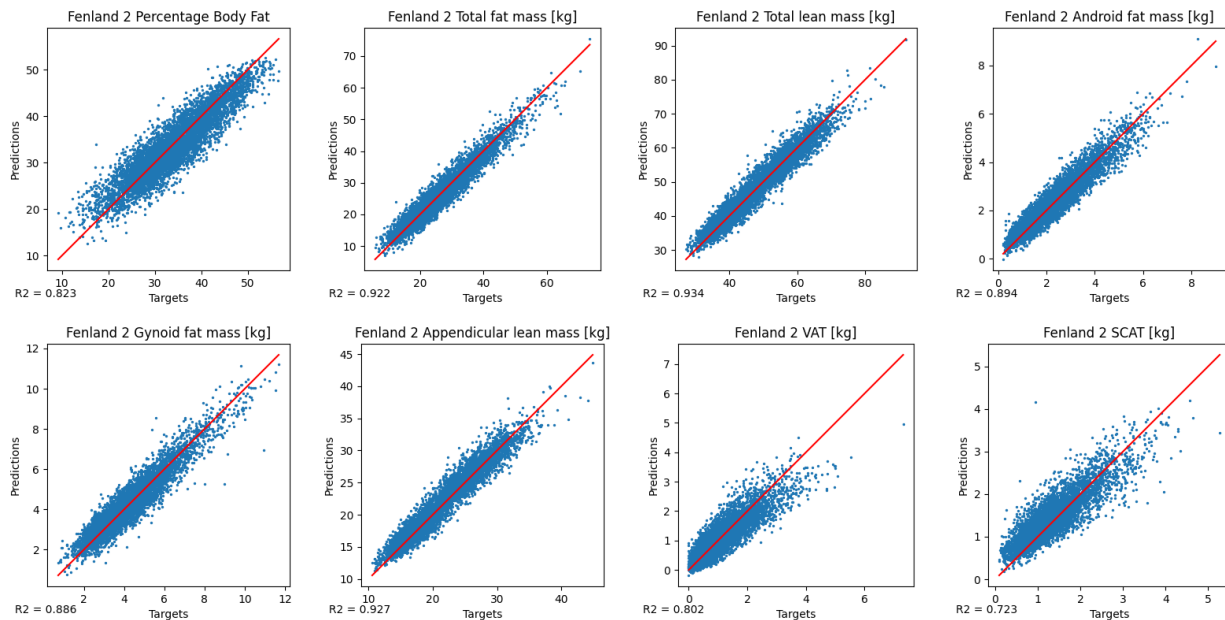


Figure 9: Scatter plots of model predictions against target values on Fenland 2. Our model is capable of predicting numerous composition metrics with high accuracy.