Title

A device-agnostic shape model for automated body composition estimates from 3D optical scans.

Journal

Medical Physics, 49(10)

Authors

Tian, Isaac
Wong, Michael
Kennedy, Samantha
et al.

Publication Date

2022-10-01

DOI

10.1002/mp.15843

Peer reviewed

HHS Public Access

Author manuscript

Med Phys. Author manuscript; available in PMC 2023 October 01.

Published in final edited form as:

Med Phys. 2022 October ; 49(10): 6395–6409. doi:10.1002/mp.15843.

A device-agnostic shape model for automated body composition estimates from 3D optical scans

Isaac Y. Tian[1], **Michael C. Wong**[2], **Samantha Kennedy**[3], **Nisa N. Kelly**[2], **Yong E. Liu**[2], **Andrea K. Garber**[4], **Steven B. Heymsfield**[3], **Brian Curless**[1], **John A. Shepherd**[2]

[1]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, Washington, USA

[2]University of Hawaii Cancer Center, University of Hawaii - Manoa, Honolulu, Hawaii, USA

[3]Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, Louisiana, USA

[4]UCSF School of Medicine, University of California - San Francisco, San Francisco, California, USA

Abstract

Background: Many predictors of morbidity caused by metabolic disease are associated with body shape. 3D optical (3DO) scanning captures body shape and has been shown to accurately and precisely predict body composition variables associated with mortality risk. 3DO is safer, less expensive, and more accessible than criterion body composition assessment methods such as dual-energy X-ray absorptiometry (DXA). However, 3DO scanning has not been standardized across manufacturers for pose, mesh resolution, and post processing methods.

Purpose: We introduce a scanner-agnostic algorithm that automatically fits a topologically consistent human mesh to 3DO scanned point clouds and predicts clinically important body metrics using a standardized body shape model. Our models transform raw scans captured by any 3DO scanner into fixed topology meshes with anatomical consistency, standardizing the outputs of 3DO scans across manufacturers and allowing for the use of common prediction models across scanning devices.

Methods: A fixed-topology body mesh template was automatically registered to 848 training scans from three different 3DO systems. Participants were between 18 and 89 years old with body mass index ranging from 14 to 52 kg/m^2. Scans were registered by first performing a coarse nearest neighbor alignment between the template and the input scan with an anatomically constrained principal component analysis (PCA) domain deformation using a device and gender specific bootstrap basis trained on 70 seed scans each. The template mesh was then optimized to fit the target with a smooth per-vertex surface-to-surface deformation. A combined unified PCA model was created from the superset of all automatically fit training scans including all three

Correspondence Isaac Y. Tian, Paul G. Allen School of Computer Science and Engineering, University of Washington, Box 352355, Seattle, WA 98195-2355, USA. iytian@cs.washington.edu.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

devices. Body composition predictions to DXA measurements were learned from the training mesh PCA coefficients using linear regression. Using this final unified model, we tested the accuracy of our body composition models on a withheld sample of 562 scans by fitting a PCA parameterized template mesh to each raw scan and predicting the expected body composition metrics from the principal components using the learned regression model.

Results: We achieved coefficients of determination (R^2) above 0.8 on all nine fat and lean predictions except female visceral fat (0.77). R^2 was as high as 0.94 (total fat and lean, trunk fat), and all root-mean-squared errors were below 3.0 kg. All predicted body composition variables were not significantly different from reference DXA measurements except for visceral fat and female trunk fat. Repeatability precision as measured by the coefficient of variation (%CV) was around 2–3x worse than DXA precision, with visceral fat %CV below 2x DXA %CV and female total fat mass at 5x.

Conclusions: Our method provides an accurate, automated, and scanner agnostic framework for standardizing 3DO scans and a low cost, radiation-free alternative to criterion radiology imaging for body composition analysis. We published a web-app version of this work at https://shapeup.shepherdresearchlab.org/3do-bodycomp-analyzer/ that accepts mesh file uploads and returns templated meshes with body composition predictions for demo purposes.

Keywords

3D scanning; body composition; dual X-ray absorptiometry; fixed topology mesh; linear regression; obesity; principal component analysis; regional composition

1 | INTRODUCTION

Metabolic syndrome is strongly correlated with the leading causes of death in the US and the world.[1] Many clinical studies have shown the importance of regional body composition as a predictor for metabolic disease risk and increased mortality even when controlling for total body variables such as weight and body mass index (BMI). Goodpaster et al.[2] showed that males in a normal BMI range with high visceral fat mass were twice as likely to have metabolic syndrome that can lead to higher risk of heart attack and stroke. Wilson et al.[3] showed that high trunk-to-leg volume ratios predict greater diabetes risk, with the highest quintile of the population 6.8 times as likely to become diabetic. Zhang et al.[4] demonstrated a correlation between abdominal obesity and death from cancer over a 16-year longitudinal study, with a 63% increase in mortality risk in the highest quintile. Although the criterion methods for measuring body composition and metabolic risk factors reside mainly in radiology facilities, many mortality predicting variables have visually observable external effects on body shape.[5]

3D optical (3DO) scanners capture external body shape and are relatively inexpensive, noninvasive tools for gathering data that can predict or prevent metabolic disease.[6] However, algorithms that estimate body composition and metabolic risk factors from 3D scans are often proprietary and unique to a particular scanning system, such as the independent black boxed methods of Fit3D, Styku, and Size Stream. Even within data from a single system, the order, placement, and number of vertices in the file vary from scan to scan, making unified

cross-compatible prediction algorithms based on surface geometry difficult to develop and verify. These limitations undermine the accessibility of 3DO scanning as a universal clinical tool due to possible inconsistent interpretations depending on the scanning system deployed.

Previous works have attempted to bridge the divide between inconsistent system data and consistent translation to medically informative statistics. Ng et al.[7] fit a 3D human mesh template to raw scans from a single scanning system using the mesh deformation methods of Allen et al.[8] This work translated unorganized raw scans into topologically and anatomically consistent 60 001 vertex meshes (60 k), which abstracted away the specific output format of the scanning device and allowed statistical methods such as principal component analysis (PCA) to be applied.

While this method in theory was not dependent on the input system, it was only benchmarked on a single system and did not scale well to large datasets due to the manual effort involved. Point correspondences between the standard template mesh and the raw scan required manual annotation on each raw scan to initialize the mesh correspondences and deformation. PCA models built on one scanning system did not generalize to scans from novel systems due to strict pose constraints inherent in the single scanning device. A scalable, device agnostic solution needs to allow for fast automatic mesh template fitting across multiple devices while showing high accuracy on body composition prediction in held-out test data.

The primary objective of this study was to develop an algorithm for automatically standardizing 3DO scans from multiple total body scanning systems captured in standard anatomical poses and predicting body composition metrics from the resulting standardized body meshes using a unified scanner agnostic model. While Ng et al.[7] directly applied the per-vertex deformation algorithm detailed in Allen et al.[8] with the assistance of labeled correspondences between template and raw scan, we first used a global deformation constrained by a bootstrap principal component basis created with a subset of manually guided template fits (see Supplementary Methods) to enforce anatomical correspondence. We refined the surface-to-surface alignment with the same per-vertex deformation but with zero manually annotated markers as part of the optimization as anatomical and topological consistency were constrained by the principal component domain. We extended the manual fitting method of Ng et al. to automatically fit a standard template to 848 training scans from Systems 1–3.

We performed a cross-sectional study on a diverse sample of convenience that received metabolic status measures and 3DO scans on four different technologies, including one unseen technology that was not used in model training. A secondary aim of this study was to quantify the test-retest precision of body composition estimates using our automatic mesh templating method against DXA.

2 | METHODS

2.1 | Study design

We constructed a parameterized 3D body shape statistical model to both perform the 3D geometric surface registration and the subsequent body composition prediction from the standardized mesh template. Our method was trained and tested on adult participants from the Shape Up! Adults Study (NIH R01 DK109008) and were recruited in the Honolulu, HI area at the University of Hawaii at Manoa (UH), in the San Francisco, CA area at the University of California, San Francisco (UCSF), and in the Baton Rouge, LA area at Pennington Biomedical Research Center (PBRC) as described in Tian et al.[9] Recruitment was stratified by age (18–40, 40–60, >60 year), ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian, and Native Hawaiian or Pacific Islander), gender and BMI (< 20, 20–25, 25–30, > 30 kg/m^2).

Participants were excluded if they could not stand unassisted for 2 min or lie supine for 10 min without movement, had metal objects in their body, or had major body-shape-altering procedures (e.g., liposuction, amputations, etc.). Female participants were excluded if pregnant or lactating. Written informed consent was obtained from each participant upon arrival, and all procedures were approved by the PBRC Institutional Review Board (IRB# 2016–053-PBRC), the UH Office of Research Compliance (CHS# 2017–01018), and the Human Research Protection Program Institutional Review Board at the UCFS (IRB# 15–18066). The study is publicly listed on ClinicalTrials.gov as ID NCT03637855.

Ground truth total and compartmental body composition measurements were defined by DXA. We acquired duplicate whole-body scans of each participant on either a Hologic Horizon/A system (UCSF) or a Discovery/A system (PBRC and UHCC) (Hologic Inc., Marlborough, MA, USA). Participants were positioned and scanned according to each manufacturer's guidelines. All scans were analyzed at UHCC by a single certified technologist using Hologic Apex version 5.6 with the National Health and Nutrition Examination Survey Body Composition Analysis calibration option disabled. DXA systems quality control was performed by monitoring the weekly values of the Hologic Whole Body Phantom. Cross-calibration was checked between sites using a whole-body phantom scanned at each site. No cross-calibration adjustments were needed.[7]

Each participant was scanned in one or more 3DO systems pending availability at each recruiting location. We used four different 3DO system manufacturers across all sites: System 1 (Fit3D Proscanner 4.x, Fit3D Inc, Redwood City, CA, USA), System 2 (Styku S100 4.1, Styku LLC, Los Angeles, CA, USA), System 3 (Size Stream SS20, Size Stream, Cary, NC, USA), and System 4 (Naked Body Scanner, Naked Labs, Redwood City, CA, USA). Scans from different systems differed slightly in pose, although all were upright with straight elbows and knees in a neutral A-pose, with elbows and knees held in maximum extension and arms and legs abducted slightly away from the midline of the body, and differed significantly in vertex count, spanning three orders of magnitude from 4000 to 400 000 points. Participants stood with arms and legs held straight and slightly away from the midline of the body, but the exact angles varied with the position of the handrails, foot marker placement (if present at all), and height of the subject. System 2 had more extreme

arm position variance as there were no fixed handrails. In the case of System 4, the restricted field of view of the optical sensor created much more variance in the pose of participants, who had to stoop or bend their limbs to fit within the constraints of the system. Participants were scanned twice to gather a test-retest precision evaluation data set.

The number of raw scans included in this study was as follows: for males, System 1: 241, System 2: 216, System 3: 116, and System 4: 59; for females, System 1: 294, System 2: 276, System 3: 135, and System 4: 73. Some participants were represented by more than one system. These systems spanned a wide range of output resolution, with some System 3 scans having as few as 4000 vertices and the largest System 1 scans having over 400 000. Scanner properties are summarized in Table 1. We standardized all scans to System 1 reference coordinates shown in Figure 1. We trained our PCA model on a subset of the first three systems and treated System 4 as a completely unseen validation method with extreme pose variation. For each training system, the participants for each gender were split into three sets: a 70-member bootstrap set and the remainder split 50/50 into train and test.

3 | RESOLVING DATA INCONSISTENCIES USING STANDARDIZED MESH TEMPLATES

A key advancement on previous work predicting body composition from unorganized 3DO scans was substituting the manual annotation of anatomical landmarks for a coarse initial deformation in principal component domain.

To fit our 60 k standardized mesh template to the scan point cloud, we iteratively minimized closest-point Euclidean distances between vertices in our template and the input scan in two phases. First, we constructed six PCA bases using manually fit[7] template meshes from the bootstrap set to initialize the automatic fitting algorithm. There were three scanning systems, and PCA spaces were separated by gender. Each bootstrap set consisted of 70 fitted meshes. We performed a global mesh deformation of our template body mesh in the dimension-reduced bootstrap principal component domain corresponding to the mesh's gender and scanning system to constrain surface smoothness and anatomical consistency as described in Tian et al.[9] and achieve a coarse shape fit. Closest point pairs were determined with a nearest-neighbor algorithm, and pairwise distances were minimized with a linear least-squares solver. This process was repeated iteratively until the difference between iterations fell below the convergence tolerance hyperparameter. Second, we refined the surface alignment between the deformed template and the scan mesh by optimizing for per-vertex 3D rigid transformations to minimize surface-to-surface distance as shown in Allen et al. to produce the final fit. This step was analogous to the mesh fitting in Ng et al. but was fully automatic with no annotated anatomical landmarks, as deformation in the principal component domain constrained the coarse fit to be topologically and anatomically consistent with a human body shape. We first fit the 420 scans from all bootstrap sets with the initial manually bootstrapped PCA models. Mathematical details are provided in the Supplementary Materials.

3.1 | Verifying anatomical and topological consistency of automatic template fits

We compared our automatic markerless body shape fits to the manually fit equivalents from Ng et al. to check for topological and anatomical consistency between the two shape fitting methods. The 60 k mesh template had 74 anatomical markers placed at anthropometric positions originally defined by the CAESAR study[10] by a trained medical professional. The original landmarks corresponded to skeletal features determined by palpation on a live subject. Our manually annotated landmarks were placed in a 3D model viewer by a trained technician to correspond to the physically palpated ones.

We can recover the 3D position of these landmarks in any shape fit with the 60k template by querying the coordinates of the neighboring vertices that define the placement of each landmark. We compared the landmarks recovered from automatic template fits to the equivalent manually clicked landmarks, using the latter as the gold standard, and compared their precision with manually annotated equivalents to the benchmarked precision of repeat manual point placement on the same scan.

3.2 | Expansion of body shape model using markerless automated fitting

We repeated building the initial six separate bootstrap PCA models but with the resulting automatically fit mesh templates and compared their body composition prediction accuracy against the manually fit baselines to ensure there was no loss of accuracy due to dropping manual annotation from the pipeline. Each model was incompatible with scans from a different system. This incompatibility was due to the strict hand and feet endpoints imposed by each system, which introduced minor pose variations but caused meshes scanned with one system to be unrepresentable in the PCA domain of another. Figure 2 shows an example of a System 2 scan fit using a System 1 PCA male model.

Our goal was to train a single PCA model that could automatically fit templates to raw scans from any system and predict accurate body composition from the standardized fit. We used the second automatically fit bootstrap model for each gender and system combination to fit all the raw scans from their respective training sets. These coarse PCA fits were then refined with surface-to-surface alignment as previously described. The refined fits from all systems were grouped into one training set per gender.

We constructed the final unified PCA model by merging the refined fits of all training data for all scanning systems into one training set per gender. The combined training data included the training set and the bootstrap set, which was a specific subset of the training data. This system-agnostic model included 391 males and 457 females. The unified model was used to fit raw scans from the held-out test set and predict their body composition.

3.3 | Predicting body composition using scanner agnostic shape space

We solved for a linear mapping between PCA basis coordinates and body composition by performing linear regression between the principal components of the automatic template fits to training scans and their corresponding criterion body composition measurements taken with DXA as described in Tian et al. Linear models for predicting body composition as a function of principal component basis weights were solved with least squares regression

between the principal components and the ground truth DXA variables. For a regression of the form **Mx** = **y**, we augmented the vector **x** with the value 1 to allow for nonzero intercepts.

We reported all prediction accuracies on test meshes that were not used in PCA model construction or regression training. Statistical significance for body composition prediction was computed with a paired two-sided *t*-test. Body composition prediction from 60k mesh templates were paired with the DXA-derived gold standard measurement. The test was successful if the null hypothesis could not be rejected, that is, the difference between DXA and our prediction was not significantly different from 0. The Bonferroni correction for the *p*-value was 0.004. $n = 182$ and 248 total test set meshes for males and females, respectively, with an additional separate 59 and 73 validation meshes from System 4. Mathematical details are written in the Supplementary Materials.

3.4 | Algorithm workflow summary

Training procedure: START: 70 male and 70 female bootstrap scans from each of Systems 1–3. (420 total)

1. Manually assign anatomical correspondences between 60 k template and bootstrap scans. Perform marker-guided surface to surface deformation to make 6 bootstrap template sets.
2. For each device AND gender combination (6 total System # + M/F combinations), make an independent 70-member bootstrap basis with PCA.
3. For each bootstrap scan, deform the 60k template to satisfy nearest neighbor alignment in its gender and scanner-specific PCA bootstrap domain. See Supplementary Materials.
4. Starting with coarse alignments from step 3, smoothly deform each mesh to register to target scan with per-vertex surface-to-surface alignment (Allen et al.)
5. Repeat step 2 with the results from step 4. Verify marker alignment with manual placements.
6. Repeat step 3–4 using the PCA model from step 5 to fit all training scans.
7. Merge all refined fits from step 6 into a single superset per gender consisting of fitted training meshes from all three scanning devices. Perform PCA on this fixed topology set to get the unified PCA basis. (1 male, 1 female)
8. Project all refined fits from step 4 onto the PCA basis from step 7 to get PC basis coordinates for each training mesh. Learn per gender linear regressions from PC coordinates to DXA measurements. (Tian et al.)

Testing procedure: For any NEW test scan (any device):

1. Fit a 60 k mesh template using PCA initialization as described in training step 3 and 4, using the PCA superset model from training step 7 as the coarse deformation prior.

2. Project the refined test mesh fit onto the unified PCA basis as in training step 8. Use regression matrix learned in training step 8 to estimate DXA measurements.

4 | RESULTS

Participant characteristics for the training and test sets are shown in Table 2. There were 1278 scans from three scanning devices randomly divided into training and testing. We reserved 70 manually fitted scans of each gender from the training set of each of the three systems to bootstrap the shape model. There was no significant difference between the means of training and test data.

4.1 | Landmark consistency between automatic and manual fits

To compare the topological consistency of our automatic fitting method to the manually guided fitting method of Ng et al.,[7] we compared the pairwise Euclidean distances between the 74 anatomical landmarks specified in the CAESAR dataset on both sets of template fits. The landmarks on the base template mesh were readily mapped to any automatically fit mesh as a function of neighboring vertices. Between topologically consistent template fits to the same raw scan, vertices and therefore landmarks should be analogous. For the 420 scans in the bootstrap set that had manually placed markers for reference, mean error was 11 mm with a standard deviation of 12 mm. By comparison, the precision of a trained analyst manually assigning markers on the same mesh for a set of 15 males and 15 females was 8 mm with a standard deviation of 7 mm.

A subset of test set meshes had manual fit equivalents, although they were not used in training. To check the anatomical correspondences for held-out test meshes, we compared the markers of 95 male and 135 female test set meshes fit with our automatic template fitting across all three scanning systems to their manually placed counterparts. Test set meshes were fit with the final unified PCA space, which was composed of automatically template fit meshes from the training sets of all three systems. The mean distance between the recovered markers and the manually placed equivalents was 21 mm, with a standard deviation of 23 mm. Figure 3 shows a visual comparison of manually placed markers to the markers recovered from the auto-templated fit.

Automatic mesh templating and standardization exhibits greater point placement error compared to repeat manual annotations, although surface-to-surface registration was visually equivalent. We show below that the slightly varied vertex distribution on a common surface did not decrease prediction accuracy of body composition prediction from body shape.

4.2 | Accuracy of markerless, unified shape models against manual baseline

Figure 4 shows automatic template fitting using the unified superset model for coarse shape alignment on the heaviest female and tallest male in the test set. A visualization of our fitting method on raw scans of a single participant scanned on four different systems is shown in Figure 5.

We recreated the same six manual bootstrap PCA models using automatically templated meshes to isolate the effect of automatic template fitting on body composition prediction

in the absence of additional training data. Figure 6 shows the difference in R^2 prediction on test data between automatically and manually fitted PCA models containing the same 70 bootstrap meshes for training. Most metrics gain predictive accuracy for both genders, with small increases in error of <0.1 for some measurements such as arm fat in females. R^2 values for the auto-templated models increased by an average of 0.054 and 0.023 for males and females, respectively, averaged across all three systems and 12 composition metrics. Although the training meshes were the same, there was an overall increase in predictive performance. This boost can be attributed to the smoothing operation followed by markerless nearestsurface smooth alignment. The resulting fits from our method had fewer artifacts that were introduced as a byproduct of the manually initialized deformation, where the initial template mesh was very misaligned with the target scan and necessitated very large nonlinear transformations.

We present the unified superset model consisting of the union of all automatically fit template meshes from the training sets of all scanning systems per gender as the benchmark model in this work. Figure 7 shows the body composition prediction R^2 difference on system-specific test scans between predictions learned from the unified superset PCA model and the 70 member manually fit bootstrap of each scanning system. Our unified superset had many more training examples than the bootstrap model (391 for males and 457 for females) but had increased noise introduced by pose variance. Our results show that the unified superset model produced good geometric fits to withheld test scans and improved on the body composition predictions made by the initial manual model despite the inclusion of training scans that exhibited differing pose.

4.3 | Body composition prediction accuracy and precision on test data from multiple systems

We automatically fit standardized mesh templates to test set scans from all three systems and used the projected PCA basis coordinates to predict body compositions from our regression matrices. Table 3 shows the R^2, root-mean-squared errors (RMSE), and p-values of the predictions on test scans using the final unified model and all available PCA components (d = 391 for males and 457 for females) compared to DXA as reference. Percent fat was calculated as fat mass/scale weight, fat mass index (FMI) was calculated as fat mass/height^2, and fat free mass index (FFMI) was calculated as lean mass/height^2. Lean mass was calculated as weight - fat mass. All masses are reported in kilograms (kg). Scans were sourced from three input systems and varied between 4000 and 400 000 vertices. All test scans were held out of PCA model training and composition regression. p-values were calculated with a paired t-test against respective DXA measurements to determine the presence of bias in our method. For 12 simultaneous measures, a $p < 0.05/12 = 0.004$ was considered significant, meaning the mean difference between our predictions, and DXA was statistically significantly different from 0. Our bias check passed for every measure except visceral fat in both genders, and in trunk lean in females. Our results were comparable to the PC-only model of Ng et al. but were reported on completely held out test data, tested on scans from two additional scanning systems, and required no manual intervention to initialize the fit. A breakdown of the results by input device is displayed in Table S1.

Table 4 shows our test-retest precision measured with the coefficient of variation (%CV) Glüer et al.[11] Same-day duplicate scans were taken of the subjects in the test set, and the fitting and predictions were repeated using the final unified model. Coefficient of variation (%CV) was defined as the ratio of the standard deviation of repeat measurements to the mean of repeat measurements averaged across all test subjects. The closer this value was to zero, the more precise our predictions for repeat scans of the same participant. Not every participant had a duplicate scan, so the test–retest pairs were less than the total test set size. We compared test–retest precision against the duplicate DXA measurements for each participant to determine precision of our method for predicting the composition of fitted scans relative to the gold standard method.

4.4 | Template fitting and body composition prediction on novel system (System 4) input

To test the generalizability of our fitting and prediction method to an unseen scanning technology, we performed automatic template fitting and body composition prediction using the unified system-agnostic model on 59 males and 73 females scanned with System 4. This device had no representation in the training set and presented an additional challenge of having the most nonconforming pose of any of the optical systems. The very limited field of view of this system necessitated many participants to hold their arms very close to their body often in bent positions, and taller individuals had to stoop either by bending at the back, waist, or knees to fit into the scanning volume. These meshes were not well conforming to the A-pose constraints specified in the meshes of the training set, in which the abduction of the arms and legs varied but all subjects stood upright with fully extended knees and elbows. R^2 and RMSE are reported in Table 5, and test–retest precision is reported in Table 6. Although the difference between our prediction and DXA was statistically significant from zero for more compositional measures relative to the devices that were included in the training data, the total body lean mass, fat mass, and percent fat predictions were not significantly different from the gold standard despite the unfavorable poses in this validation set.

Our method worked well for most scans on this validation system but did not align properly to some scans with larger pose differences, such as excessively bent elbows. These scans were excluded from our results as they violated the parameters of the model. Our method can be used to fit a standardized templated to scans from any system that captures a human in an A-pose, which is any pose with fully extended elbows and knees with arms and legs abducted around 30 degrees from the midline. Some variation in limb abduction at the shoulders and hips is accounted for by the model, but large body part rotations create nonlinearities not representable by PCA.

5 | DISCUSSION

In this work, we developed a scanning system agnostic algorithm for standardizing unorganized 3DO body scans with vertex counts spanning three orders of magnitude and slight variations in pose. We constructed a PCA model using training data from three different systems and validated our shape fitting and body composition prediction accuracy

on held out scans from all three systems and an additional fourth unseen system. Our resulting prediction models trained on automatically fit training set scans predicted total and regional body fat with R^2 and RMSE comparable to the principal component only model from Ng et al.[7] Our model was more accurate in many cases even on the validation system. For example, R^2 for total fat mass for System 4 was 0.94 and 0.96 for males and females, respectively, compared to 0.88 and 0.93 in Ng et al. Our results are even stronger when compared to previous work as our fits, and predictions were performed on test scans that were not included in the PCA training data. In Ng et al., five-fold cross validation was performed to train the linear model mapping principal components to body composition, but the PCA domain included all the available data and thus contained no withheld data for blinded validation. Furthermore, our scans were sourced from four different systems and generalized to a system that was not represented by training data with no manual initialization. We demonstrated that our automatic template fitting algorithm can generalize to inputs from novel scanning systems exhibiting small variations in pose. Inclusion of these system-agnostic template fits into a new expanded PCA domain is likely to further increase the predictive accuracy of our regression models.

Our retest precision error on the test set was two to four times that of the criterion DXA scans. This was a slightly less precise result than Ng et al. but was reported on a greater number of subjects (as opposed to just 119 of each gender) scanned on three different systems (as opposed to one). Furthermore, all test scans were held out of the PCA construction while Ng et al. made no such distinction. Although our precision is trailing that of criterion radiology, this can be mitigated by averaging predictions from repeat scans. The least significant change (LSC)[12] between the average of multiple measurements sampled at baseline and follow-up is defined as:

$$\mathrm{LSC} = Z\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

for a Z defined by the two sided 95% confidence interval z-score and precision error σ, with n_1 measurements at baseline and n_2 measurements at follow-up. A difference greater than this value means there is 95% confidence a true change in body composition has occurred, and a lower value implies greater resolution of change over time. Z is constant at 1.96, and σ is inherent to the measurement method, but we can set $n_1 = n_1 = 9$ scans at baseline and follow-up to drop the LSC by a factor of 3. This would account for the difference in precision error between our method and a single DXA scan on most metrics. As 3DO scanners take a minute or less to complete and can be repeated without threat of radiation injury, collecting nine scans is not unreasonably burdensome relative to the gold standard radiology. As we have shown the accuracy of our method to increase with larger training sets, the precision of our method could potentially be increased with additional data.

Our automated markerless fitting and prediction method generalized well to scans that were adjacent to an A-pose. We tested incorporating additional training scans in a T-pose stance, with arms held out parallel to the ground and legs straight with feet together, to determine how robust our method was to scans with more extreme pose variation. Although we were able to achieve good geometric fits to test scans in both the A and T pose, the

predictive accuracy of the regression models decreased drastically. Such huge differences in the limb positions were not correlated with body composition but had major effects on the PCA shape coordinates of the mesh. We decided against recommending inclusion of this degree of extreme pose variability in the training set for the results of this paper as the variance introduced by the extreme pose difference created excessive amounts of noise in the prediction regression. An example of a T-pose fit using our model is shown in Figure 8.

We demonstrated that our method was able to fit to a range of A-poses if there was representation of the target pose in the training data. However, these pose differences were modeled as linear interpolations of the limbs between example meshes in the training data exhibiting pose variation rather than rotational transformations at joints. A body mesh that separately models pose and shape using a jointed skeleton, and skinning function may provide more accurate fits and predictions for scans that are captured in poses that are further from anatomical neutral, that is, the hands-up pose in an airport system. Such a model may also produce increased prediction accuracy for body composition even on A-pose scans, as the vertex deformations caused by small differences in limb position are not discriminated from differences in anatomical body shape and create substantial noise in learning the relationship between shape and composition. Although prior work such as Skinned Multi Person Linear (SMPL)[13] exists in the field of human pose and shape estimation using skinned models, these methods were not created with clinical application in mind. SMPL learned a linear body shape PCA from the CAESAR database and decoupled pose from the model by standardizing all fitted templates to a T-pose with parallel arms and vertical legs. This was made possible with an animated model containing a skeleton with vertices mapped to a hierarchical tree (known as a skinned model). However, they did not collect medical measurements along with shape and pose data for their model construction. Subsequent work mainly sought to recreate visually plausible reconstructions of human shapes from “in the wild” photos[14] and lack correlated clinical variables. Future work may merge the flexibility of a skinned posable model with our clinical data.

While our work is not the first to propose templated encoding spaces for human body shapes,[15] our implementation was the first to include associated body composition data with 3DO scanning and also the first to demonstrate a robust pipeline that works with limited pose variation across different scanning devices. Due to the logistical challenges of collecting multiperson data, CAESAR[10] remains the largest fixed pose, multishape dataset almost 20 years after its initial publication. Many recent works on human body encoding focus on limited person, multipose shape models,[16] and do not address the medical implications of body shape. Future work could use our method to combine the scans of CAESAR and Shape Up! into a singular multi-person shape model and learn the correlations to body composition on the subset of training data with associated DXA measurements using more robust shape descriptors learned on the entire large dataset.

Certain fine details in 3D scans are not useful in predicting body composition, such as the position and shape of the hands and fingers or head shape deformation due to large hair volume. Further work could replace these regions with smooth surfaces, effectively eliminating the variance in shape caused by these uncorrelated details. Such a reduction in noise could potentially increase the accuracy of our prediction models.

Our algorithms predict analogs for mortality risk due to metabolic disease from 3D body shape. The relationship between scalar metrics such as compartmental body fat masses and disease has been previously studied in the clinical setting using radiology. Our method allows the use of 3DO scanning as an alternative to traditional medical imaging for predicting these measurements. It is possible that there are direct relationships between 3D body shape and disease risk latent in the templated data that can be recovered from the parameters of the principal component domain or through deep-learning methods on the mesh coordinates directly. The standardization of 3D scans allows for the application of these methods, but we cannot yet form such conclusions without a longitudinal dataset that directly associates body shape with mortality from metabolic disease. Our algorithm may be able to recover higher dimensional relationships between body shape and mortality risk in the future, bypassing scalar analogs such as prediction of visceral fat mass, when such data becomes available.

Our prediction model was determined with a least-squares linear regression between the principal component coordinates of our training data and their associated DXA composition metrics. Our method worked well in Tian et al.[9] and was the most conservative model for relating standardized 60 k mesh templates to body composition. A better performing prediction model with nonlinear terms, such as one produced by a graph convolution network in Bouritsas et al.,[17] may offer better predictions from either the dimension reduced PC coordinates or the 60k mesh vertices themselves. As these models required thousands of training examples, we relied on linear models that could learn good correlations with only a few hundred scans. Creating larger standardized body mesh databases using our method with future data collection may facilitate work in this direction.

6 | CONCLUSIONS

At the conclusion of this study, we automatically fit a common 60 k vertex body template to 1410 raw scans from four different scanning technologies with an automated PCA initialized two-stage deformation process.

We built a device agnostic-unified PCA model from 848 training scans and learned a regression from projected PCA basis coordinates to DXA body composition measurements. We used the same fitting procedure with this device agnostic model to automatically template 562 held-out test scans and derive body composition predictions from PCA coordinates. Our model predicted body composition metrics with accuracy comparable to or better than previous models built with manually targeted mesh fitting, achieving 0.8 or better R^2 for all fat and lean mass predictions on test scans except female visceral fat, with all RMSEs below 3.0 kg. Processing data at this scale by hand would be prohibitive in time and cost. Furthermore, our results were reported on held-out test data, which further strengthens our results relative to previous work published on training accuracy only. Our work seeks to make 3DO scanning an accurate, automated, and device agnostic tool for body shape modeling and composition analysis with many potentials for clinical and diagnostic application, such as serial monitoring of changes in body weight and risk for chronic disease.

ACKNOWLEDGMENTS

The authors thank laboratory assistant William M. Nguyen for support in software pipeline programming and results validation during this project, which made possible the processing of over 1000 scan fits and composition predictions. Aimee Bowen contributed as the principal web engineer and worked closely with the lead author to develop a public facing web-GUI implementation of this work, available at https://shapeup.shepherdresearchlab.org/3do-bodycomp-analyzer/. SBH: National Institutes of Health NORC Center Grants (P30DK072476, Pennington/Louisiana and P30DK040561, Harvard); JAS: National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (R01DK109008 and R01DK111698).

Funding information

National Institutes of Health NORC Center, Grant/Award Numbers: P30DK072476, P30DK040561; National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Grant/Award Numbers: R01DK109008, R01DK111698

DATA AVAILABILITY STATEMENT

The data underlying this study cannot be made publicly available because the data contain patient identifying information. Data are available from the Shape Up! studies for researchers who meet the criteria for access to confidential data. For details and to request an application, please contact the John Shepherd (johnshep@hawaii.edu) or visit https://www.shapeup.shepherdresearchlab.org.

Abbreviations:

3DO	3-dimensional optical
60k	60 001 vertex mesh template
BMI	body mass index
DXA	dual X-ray absorptiometry
PCA	principal component analysis

REFERENCES

1. Hui WS, Liu Z, Ho SC. Metabolic syndrome and all-cause mortality: a meta-analysis of prospective cohort studies. Eur J Epidemiol.2010;25:375–384 10.1007/s10654-010-9459-z [PubMed: 20425137]
2. Goodpaster BH, Krishnaswami S, Harris TB, et al. Obesity, regional body fat distribution, and the metabolic syndrome in older men and women. Arch Intern Med. 2005;165(7):777. 10.1001/archinte.165.7.777 [PubMed: 15824297]
3. Wilson JP, Kanaya AM, Fan B, Shepherd JA. Ratio of trunk to leg volume as a new body shape metric for diabetes and mortality. PLoS One. 2013;8(7):e68716. 10.1371/journal.pone.0068716 [PubMed: 23874736]
4. Zhang C, Rexrode KM, van Dam RM, Li TY, Hu FB. Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality. Circulation. 2008;117(13):1658–1667. 10.1161/circulationaha.107.739714 [PubMed: 18362231]
5. Ng BK, Hinton BJ, Fan B, Kanaya AM, Shepherd JA. Clinical anthropometrics and body composition from 3d whole-body surface scans. Eur J Clin Nutr. 2016;70(11):1265–1270. 10.1038/ejcn.2016.109 [PubMed: 27329614]

6. Wong MC, Ng BK, Kennedy SF, et al. Children and adolescents' anthropometrics body composition from 3-d optical surface scans. Obesity. 2019;27(11):1738–1749. 10.1002/oby.22637 [PubMed: 31689009]
7. Ng BK, Sommer MJ, Wong MC, et al. Detailed 3-dimensional body shape Features PREDICT body composition, blood metabolites, and FUNCTIONAL strength: the SHAPE Up! studies. Am J Clin Nutr. 2019;110(6):1316–1326. 10.1093/ajcn/nqz218 [PubMed: 31553429]
8. Allen B, Curless B, Popović Z. The space of human body shapes: reconstruction and parameterization from range scans. ACM Trans Graphics. 2003;22:587–594. 10.1145/1201775.882311
9. Tian IY, Ng BK, Wong MC, et al. Predicting 3d body shape and body composition from conventional 2d photography. Med Phys. 2020;47(12):6232–6245. 10.1002/mp.14492 [PubMed: 32978970]
10. Robinette KM, Blackwell S, Daanen H, Boehmer M, Fleming S. Civilian American and European Surface Anthropometry Resource (CAESAR), Final Report. Volume 1. Summary. SAE International. 2002. 10.21236/ada406704
11. Glüer C-C, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporosis Int. 1995;5(4):262–270. 10.1007/bf01774016
12. Shepherd JA, Lu Y. A generalized least significant change for individuals measured on different dxa systems. J Clin Densitom. 2007;10(3):249–258. 10.1016/j.jocd.2007.05.002 [PubMed: 17616413]
13. Loper M, Mahmood N, Romero J, Pons-Moll G, Black MJ. SMPL: a skinned multi-person linear model. ACM Trans Graph. 2015;34(6):1–16. 10.1145/2816795.2818013
14. Bogo F, Kanazawa A, Lassner C, Gehler P, Romero J, Black MJ. Keep it SMPL: automatic estimation of 3d Human pose and shape from a single image. Comput Vision ECCV. 2016;2016:561–578. 10.1007/978-3-319-46454-1_34
15. Pishchulin L, Wuhrer S, Helten T, Theobalt C, Schiele B. Building statistical shape spaces for 3d human modeling. Pattern Recognit. 2017;67:276–286. 10.1016/j.patcog.2017.02.018
16. Zhou Y, Wu C, Li Z, et al. Fully convolutional mesh autoencoder using efficient spatially varying kernels. 2020 arXiv:2006.04325v2. Accessed May 10, 2022.
17. Bouritsas G, Bokhnyak S, Ploumpis S, Zafeiriou S, Bronstein M. Neural 3D morphable models: spiral convolutional networks for 3d shape representation learning and generation. Paper presented at: 2019 IEEE/CVF International Conference on Computer Vision (ICCV); October 27—November 02, 2019; Seoul, Korea (South). 10.1109/iccv.2019.00731

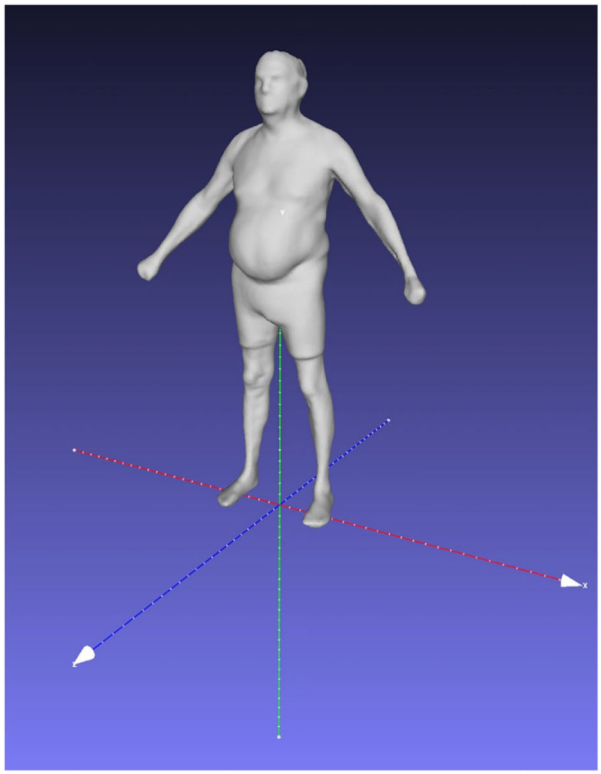

FIGURE 1. Diagram of rotation/translation relative to the origin and standard basis.
Scale is metric. This reference frame was the factory default for System 1, but in principle any common transformation could be used if all scans aligned at the feet

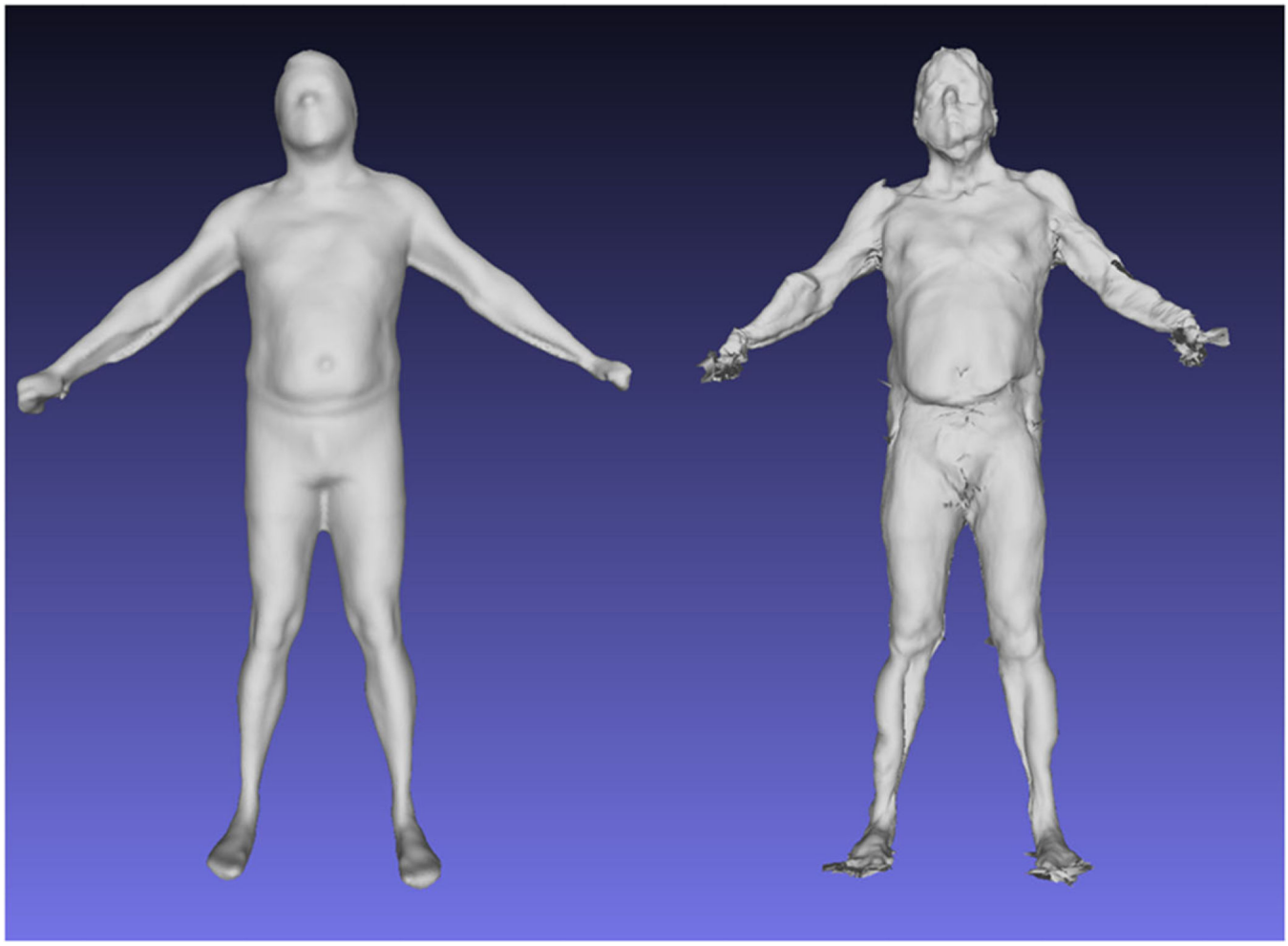

FIGURE 2. Example of Failure of Single Device Model.

Raw System 2 scan on the left, and fit attempt with System 1 only model on the right. The pose difference between the two systems (System 2 has no fixed hand position) makes models trained on a single system nongeneralizable to other systems. This demonstrates the need for a unified model that can interpolate the pose differences between different systems

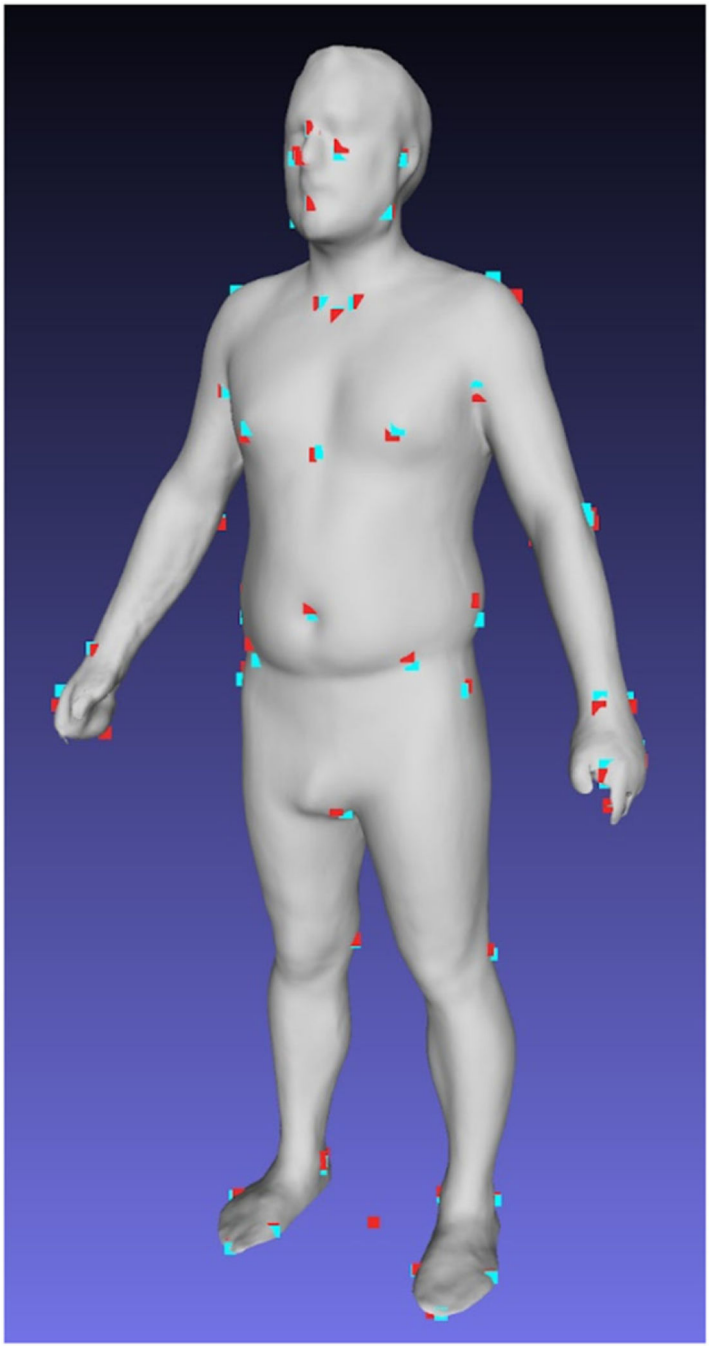

FIGURE 3. Topological consistency of auto template mesh.

Auto fit (red) and manually placed (blue) anthropometric landmarks on a fitted test set mesh. Mean difference was 21 mm for 74 landmarks across 230 meshes. The red dot between the middle of the feet is the origin of the coordinate system

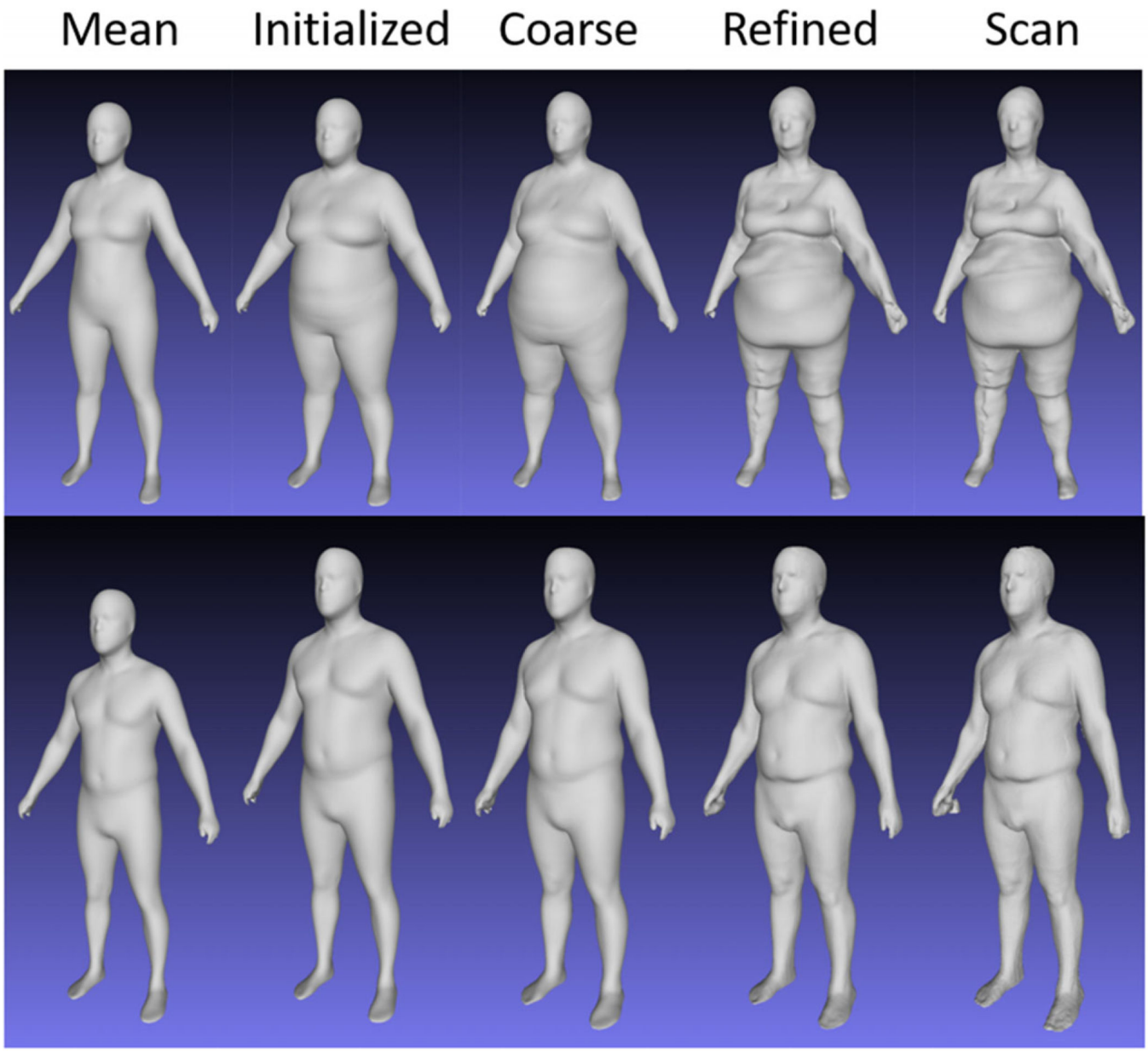

FIGURE 4. Shape fitting progression.

Top: From left to right: The mean shape, the initialization shape, the coarse fit, the final fit, and the raw input from System 1. This example represents one of the worst-case scenarios as this individual was the heaviest female in our test set at 163.4 kg and is one of the farthest shapes from the mean. Bottom: Same progression as above but using the tallest male in the test set, at 190.8 cm. Note the large size increase after initializing the body shape with the known height and weight of the participant

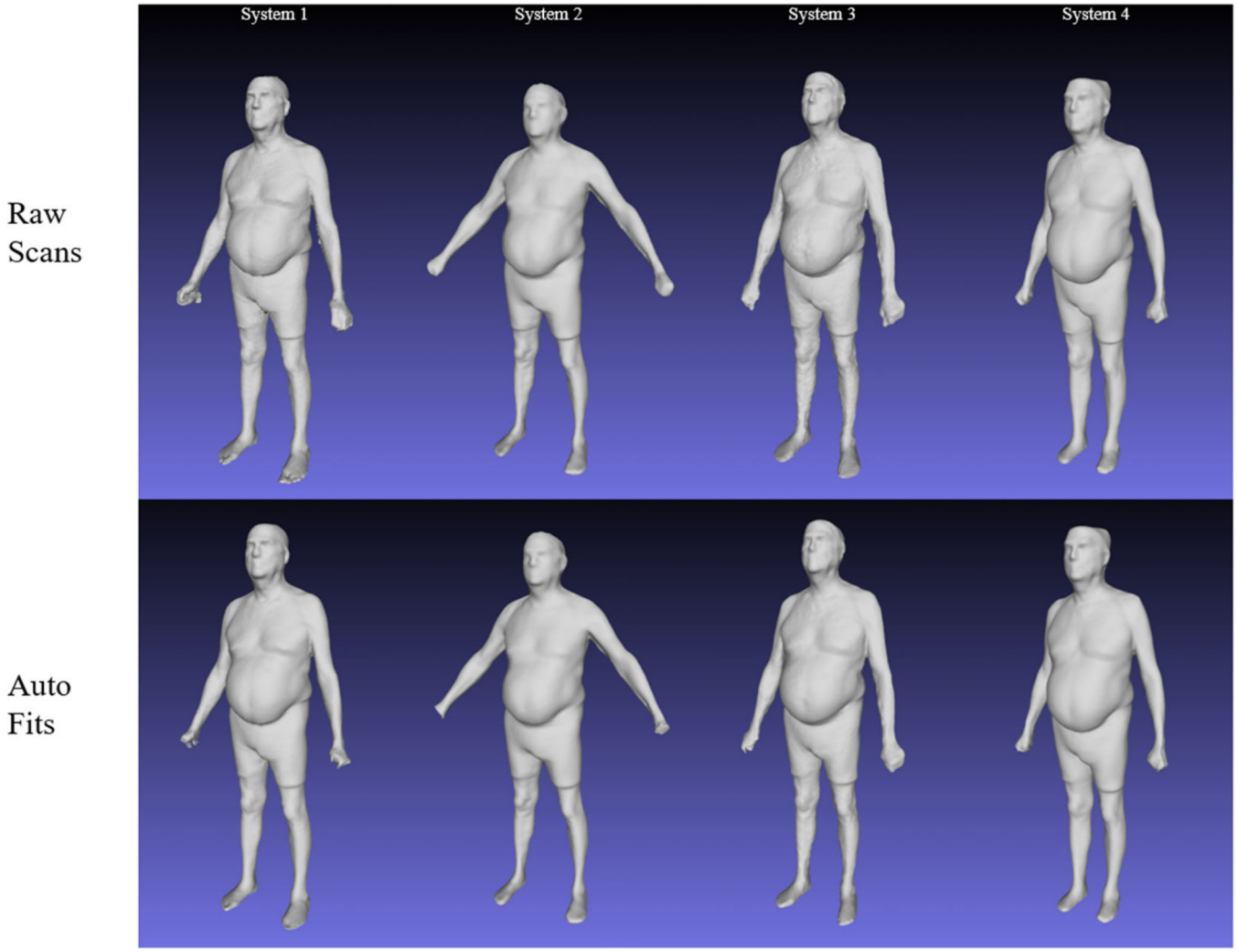

FIGURE 5. Examples of auto template fits on a single participant.

One participant was represented in the test set for all four scanning systems. Raw scan input on the top and our automatic mesh fit on the bottom. Although visually similar, the raw scan had as high as 400 000 (System 1) unorganized vertices and as low as 25 000 (System 3). We fit a 60 k anatomically consistent mesh to each scan using d = 391 principal components

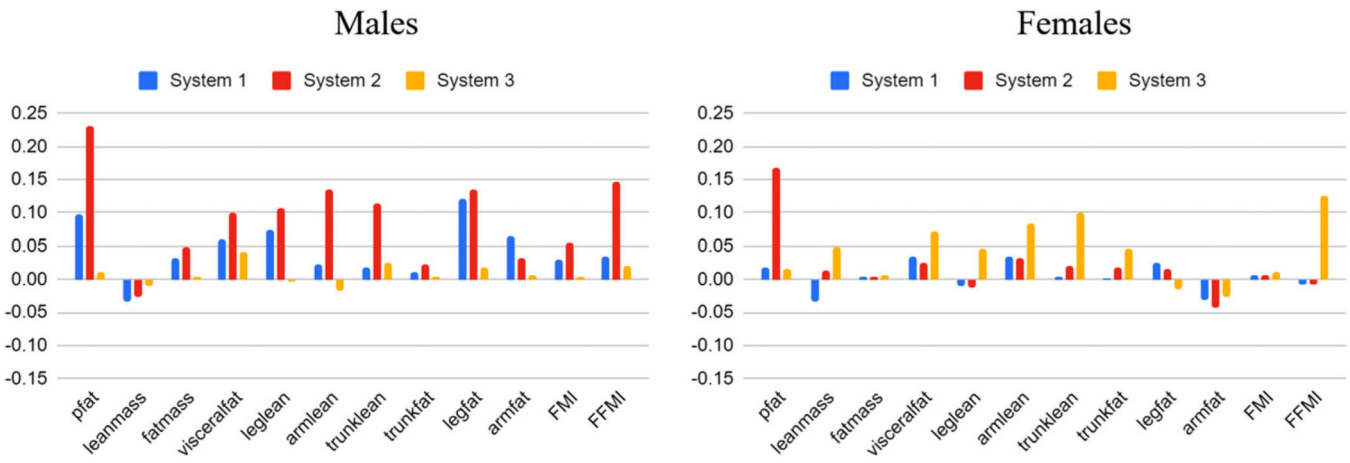

FIGURE 6. R^2 difference in body composition predictions between automatic and manual fit bootstrap models.

R^2 differences between predictions from automatically and manually fit meshes are shown on held-out test set. Training set members were the same 70 scans for each system and gender combination for both automatic and manual fitting. The only variable changed was the method of template fitting to the raw scans. R^2 increased by 0.05 and 0.02 on average across males and females, respectively, indicating substituting markerless automatic fitting for manual fitting did not cause enough topological inconsistency to impact the resulting body composition regressions

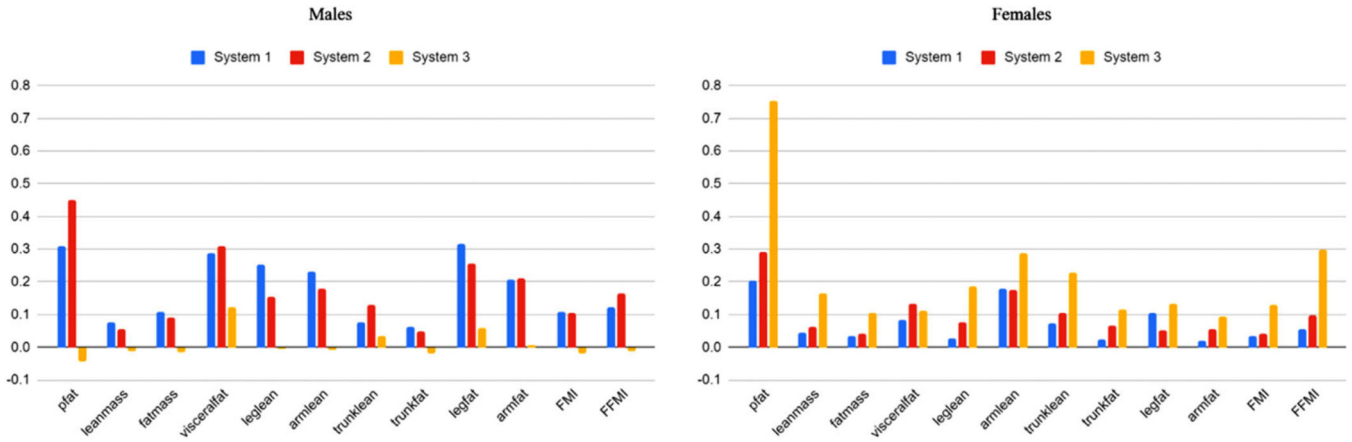

FIGURE 7. R^2 difference in body composition predictions between final and initial model.

R^2 differences are shown on held-out test set, which was the same for both the final and initial models. Initial model was composed of 70 manually guided template fits per system and gender. Each system and gender combination was a separate principal component analysis (PCA) model, resulting in six models. Models could only be tested on the system they were trained on. Final models were unified across three scanning systems but split per gender. Final male and female models had 391 and 457 training scans, respectively, and all were automatically template fit with our method. The same model could be applied to test scans from all systems. Overall, prediction accuracy increased on test data despite using no manual fits in our final model and injecting noise in the body shape model due to including pose variance from multiple scanning systems in our unified model. Per-system composition prediction metrics are shown in Table S1

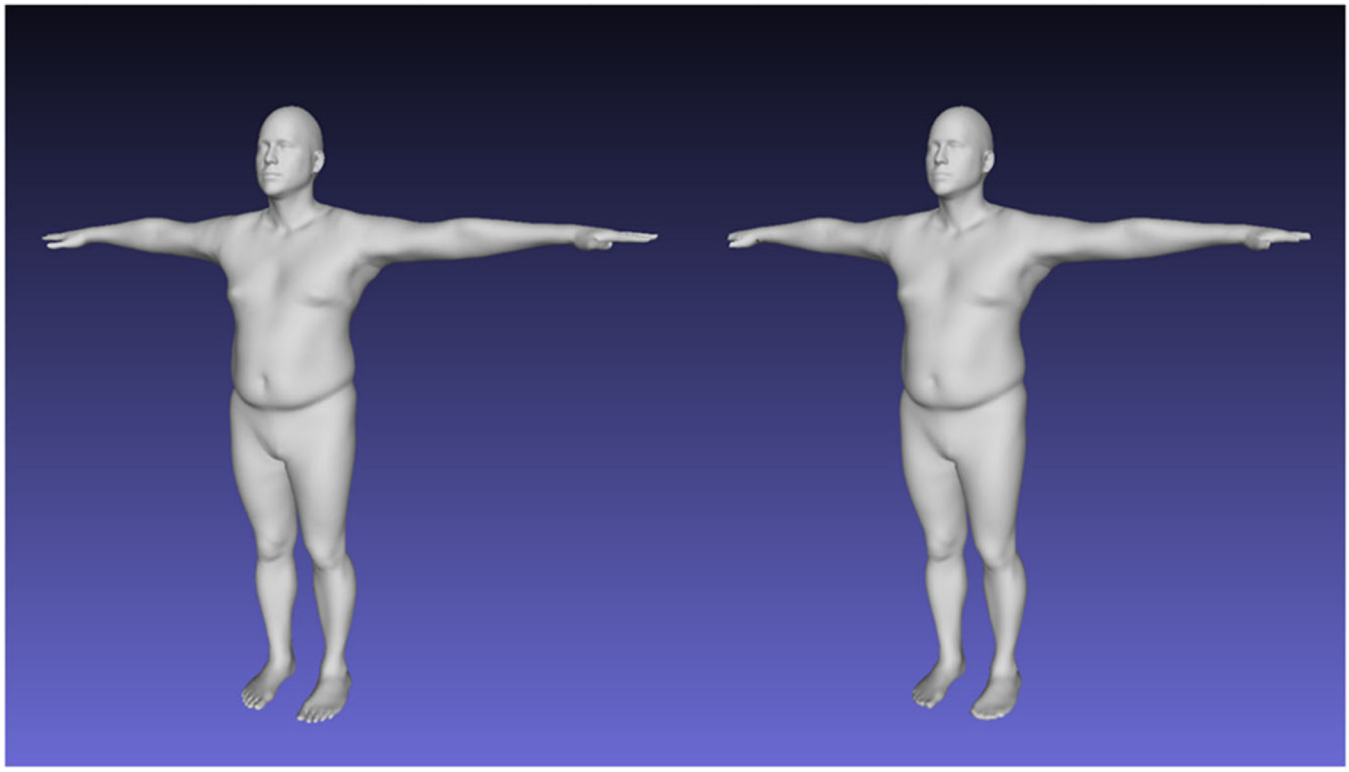

FIGURE 8. Robustness of fitting algorithm to extreme pose variation.
T-pose example, input scan on the left with 110 k vertices and our 60 k fit on the right

TABLE 1

File properties and scan times of 3DO scanners in study

	Vertices	File size	Scan time
Fit3D Proscanner (System 1)	300–500 k	40–60 MB	40 s
Styku S100 (System 2)	30–60 k	4–7 MB	20 s
Size Stream SS20 (System 3)	4–25 k	0.5–3 MB	3 s
Naked Body Scanner (System 4)	100 k	12 MB	5 s

TABLE 2

Population statistics for training and test scans. Plus-minus values are standard deviation

	Male					
	Train (*N* = 391)			**Test (*N* = 241)**		
	Mean ± SD	**Min**	**Max**	**Mean ± SD**	**Min**	**Max**
Age (Years)	44.92 ± 16.06	18	79	44.07 ± 16.15	18	79
Height (m)	1.76 ± 0.08	1.51	2.02	1.75 ± 0.07	1.55	1.91
Mass (kg)	88.42 ± 21.34	40.74	172.4	84.25 ± 18.91	40.77	135.58
BMI	28.44 ± 6.21	16.52	52.55	27.44 ± 5.55	16.96	45.82
Percent fat	22.66 ± 6.86	9.03	47.69	22.07 ± 6.71	9.03	38.58
Lean mass (kg)	67.48 ± 12.99	33.95	107.82	64.87 ± 11.64	33.95	93.1
Fat mass (kg)	20.94 ± 10.67	5.01	66.48	19.39 ± 9.47	5.13	45.94
Visceral fat (kg)	0.49 ± 0.27	0.16	1.64	0.5 ± 0.32	0.16	1.64
Leg lean (kg)	10.97 ± 2.28	5.43	18.95	10.48 ± 1.99	5.43	14.74
Arm lean (kg)	4.46 ± 1.05	2.05	8.33	4.26 ± 0.98	2.05	7.38
Trunk lean (kg)	32.36 ± 6.4	15.95	51.16	31.24 ± 5,82	15.95	48.02
Trunk fat (kg)	10.5 ± 6.15	1.76	34.12	9.72 ± 5.71	1.97	26.25
Leg fat (kg)	3.41 ± 1.71	0.89	11.86	3.12 ± 1.44	0.85	9.02
Arm fat (kg)	1.25 ± 0.68	0.29	4.34	1.15 ± 0.61	0.29	3.72
FMI	6.74 ± 3.37	1.68	20.78	6.32 ± 3.01	1.91	15.49
FFMI	21.72 ± 3.53	14.18	35.65	21.16 ± 3.23	14.18	30.72
	Female					
	Train (*N* = 457)			**Test (*N* = 321)**		
	Mean ± SD	**Min**	**Max**	**Mean ± SD**	**Min**	**Max**
Age (Years)	46.24 ± 16.13	18	89	47.53 ± 16.71	18	89
Height (m)	1.62 ± 0.07	1.44	1.80	1.61 ± 0.07 *	1.44	1.76
Mass (kg)	72.43 ± 20.93	35.44	153.05	69.39 ± 19.60	35.44	153.05
BMI	27.48 ± 7.65	14.16	51.86	26.81 ± 7.05	14.16	51.86

Med Phys. Author manuscript; available in PMC 2023 October 01.

	Male					
	Train (*N* = 391)			Test (*N* = 241)		
	Mean ± SD	Min	Max	Mean ± SD	Min	Max
Percent fat	34.06 ± 7.88	12.63	53.3	33.78 ± 7.44	17.18	53.3
Lean mass (kg)	46.49 ± 9.48	28.56	80.37	44.91 ± 9.42	28.56	80.37
Fat mass (kg)	25.85 ± 12.72	6.3	72.68	24.39 ± 11.38	6.88	72.68
Visceral fat (kg)	0.45 ± 0.31	0.06	1.37	0.43 ± 0.3	0.05	1.22
Leg lean (kg)	7.48 ± 1.73	4.25	14.03	7.21 ± 1.78	4.42	13.19
Arm lean (kg)	2.42 ± 0.57	1.31	4.42	2.34 ± 0.58	1.44	4.63
Trunk lean (kg)	23.13 ± 4.91	13.71	41.59	22.26 ± 4.72	13.71	41.14
Trunk fat (kg)	11.94 ± 6.75	2.37	35.6	11.27 ± 6.21	2.48	35.6
Leg fat (kg)	4.81 ± 2.23	1.23	12.92	4.54 ± 1.98	1.23	12.34
Arm fat (kg)	1.67 ± 0.99	0.28	6.08	1.55 ± 0.84	0.4	6.08
FMI	9.82 ± 4.79	2.01	26.57	9.44 ± 4.24	2.75	24.68
FFMI	17.6 ± 3.28	11.43	29	17.33 ± 3.16	10.93	28.88

Note: *p*-value of 0.004 was considered significant after Bonferroni correction. All metrics were not significantly different between test and train except for female height, denoted by the asterisk, which is not a predicted measurement. Body composition measurements were taken from DXA.

Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat free mass index.

TABLE 3

Body composition prediction metrics on test set. R^2, root-mean-squared errors (RMSE), and paired *p*-value with dual-energy X-ray absorptiometry (DXA) measurements for all test set fits using all available principal component dimensions (391 and 457 for males and females)

	Male, *d* = 391, *n* = 182		**Female, *d* = 457, *n* = 248**	
	R^2 (RMSE)	***p*-Value**	**R^2 (RMSE)**	***p*-Value**
Percent fat	0.77 **(3.24)**	0.32	0.68 **(4.22)**	0.27
Lean mass (kg)	0.95 **(2.57)**	0.23	0.91 **(2.84)**	0.64
Fat mass (kg)	0.93 **(2.57)**	0.23	0.94 **(2.84)**	0.64
Visc. fat (kg)	0.86 **(0.12)**	**0.0***	0.77 **(0.14)**	**0.0***
Leg lean (kg)	0.89 **(0.68)**	0.04	0.91 **(0.55)**	0.01
Arm lean (kg)	0.81 **(0.42)**	0.05	0.84 **(0.23)**	0.35
Trunk lean (kg)	0.91 **(1.73)**	0.29	0.88 **(1.64)**	**0.002***
Trunk fat (kg)	0.94 **(1.45)**	0.84	0.93 **(1.67)**	0.33
Leg fat (kg)	0.85 **(0.56)**	0.04	0.92 **(0.56)**	0.58
Arm fat (kg)	0.88 **(0.22)**	0.38	0.89 **(0.28)**	0.83
FMI	0.92 **(0.85)**	0.19	0.93 **(1.12)**	0.62
FFMI	0.93 (**0.85)**	0.19	0.88 **(1.12)**	0.62

Note: One PCA model was used to fit body shape and predict composition per gender, while input scans came from three independent manufacturers. Only visceral fat for both genders and female trunk lean were significantly different from DXA measurements in a paired *t*-test. R^2 and RMSE values were comparable to the manually guided methods of (2) but were reported on held out data rather than scans that were in the PCA domain.

TABLE 4

Test-retest precision of test set scans

	This work		DXA	
	Male *n* = 146%CV (RMSE)	Female *n* = 208%CV (RMSE)	Male *n* = 143%CV (RMSE)	Female *n* = 205%CV (RMSE)
Percent fat	**(1.90)**	**(2.91)**	**(0.50)**	**(0.48)**
Lean mass (kg)	1.22 **(1.55)**	2.09 **(1.87)**	0.38 **(0.49)**	0.45 **(0.40)**
Fat mass (kg)	4.04 **(1.55)**	3.92 **(1.87)**	1.24 **(0.47)**	0.77 **(0.37)**
Visc. fat (kg)	8.83 **(0.08)**	12.25 **(0.09)**	4.79 **(0.05)**	6.3 **(0.05)**
Leg lean (kg)	2.23 **(0.46)**	2.19 **(0.32)**	0.86 **(0.18)**	0.82 **(0.12)**
Arm lean (kg)	2.8 **(0.23)**	3.64 **(0.17)**	1.05 **(0.09)**	1.51 **(0.07)**
Trunk lean (kg)	1.86 **(1.14)**	1.99 **(0.90)**	0.75 **(0.46)**	0.78 **(0.35)**
Trunk fat (kg)	4.53 **(0.86)**	4.78 **(1.04)**	2.06 **(0.39)**	1.6 **(0.35)**
Leg fat (kg)	5.03 **(0.32)**	3.89 **(0.35)**	1.88 **(0.11)**	1.08 **(0.10)**
Arm fat (kg)	6.06 **(0.14)**	5.77 **(0.18)**	2.32 **(0.05)**	2.3 **(0.07)**
FMI	3.98 **(0.50)**	3.94 **(0.73)**	1.26 **(0.16)**	0.76 **(0.14)**
FFMI	1.20 **(0.50)**	2.11 **(0.73)**	0.38 **(0.16)**	0.44 **(0.15)**

Note: Coefficient of variation (%CV) and RMSE on test set compared to repeat DXA scans of same participants. %CV was comparable to (4) and around 2–4 times the magnitude of corresponding DXA precision metrics.

Abbreviations: DXA, dual-energy X-ray absorptiometry; FMI, fat mass index; FFMI, fat free mass index.

TABLE 5

Body composition prediction on novel system (System 4) input

	Male, *d* = 391, *n* = 59		**Female, *d* = 457, *n* = 73**	
	***R*^{2} (RMSE)**	***p*-Value**	***R*^{2} (RMSE)**	***p*-Value**
Percent fat	0.85 **(2.93)**	0.95	0.72 **(4.13)**	0.46
Lean mass (kg)	0.96 **(2.41)**	0.56	0.93 **(2.56)**	0.22
Fat mass (kg)	0.94 **(2.41)**	0.56	0.96 **(2.56)**	0.22
Visc. fat (kg)	0.63 **(0.17)**	**0***	0.85 **(0.13)**	0.015
Leg lean (kg)	0.69 **(1.09)**	**0***	0.77 **(0.85)**	**0***
Arm lean (kg)	0.15 **(0.83)**	**0***	0.82 **(0.24)**	0.006
Trunk lean (kg)	0.72 **(3.06)**	**0***	0.85 **(2.01)**	**0***
Trunk fat (kg)	0.94 **(1.43)**	0.28	0.96 **(1.44)**	**0***
Leg fat (kg)	0.88 **(0.52)**	0.09	0.86 **(0.75)**	**0***
Arm fat (kg)	0.83 **(0.25)**	**0***	0.93 **(0.26)**	0.66
FMI	0.94 **(0.82)**	0.64	0.96 **(0.98)**	0.26
FFMI	0.93 **(0.82)**	0.64	0.91 **(0.98)**	0.26

Note: R^2, RMSE, and *p*-value of System 4 test set. No scans from this device were included in any training data, and participants scanned with this system often had bent limbs or hunched postures due to a narrower field of view for the sensor. Total body lean mass, fat mass, and percent fat were still not significantly different from DXA measurements.

Abbreviations: FMI, fat mass index; FFMI, fat free mass index.

TABLE 6

Test-retest precision on novel system (System 4) compared to DXA. %CV and RMSE for System 4

	This work		DXA	
	Male *n* = 41%CV (RMSE)	**Female *n* = 58%CV (RMSE)**	**Male *n* = 41%CV (RMSE)**	**Female *n* = 58%CV (RMSE)**
Percent fat	**(2.59)**	**(3.73)**	**(0.55)**	**(0.55)**
Lean mass (kg)	1.40 **(1.86)**	2.46 **(2.26)**	0.41 **(0.54)**	0.49 **(0.46)**
Fat Mass (kg)	4.69 **(1.86)**	4.34 **(2.26)**	1.43 **(0.55)**	0.83 **(0.43)**
Visc. fat (kg)	13.57 **(0.11)**	14.23 **(0.12)**	3.80 **(0.04)**	6.65 **(0.06)**
Leg lean (kg)	1.91 **(0.38)**	3.87 **(0.53)**	0.86 **(0.18)**	0.9 **(0.13)**
Arm lean (kg)	3.21 **(0.23)**	5.65 **(0.27)**	1.25 **(0.11)**	1.38 **(0.07)**
Trunk lean (kg)	2.55 **(1.51)**	2.99 **(1.35)**	0.74 **(0.48)**	0.96 **(0.45)**
Trunk fat (kg)	6.16 **(1.20)**	4.79 **(1.12)**	2.21 **(0.44)**	1.73 **(0.43)**
Leg fat (kg)	4.75 **(0.31)**	5.09 **(0.52)**	1.80 **(0.11)**	0.92 **(0.09)**
Arm fat (kg)	7.13 **(0.18)**	5.10 **(0.17)**	3.12 **(0.07)**	2.58 **(0.09)**
FMI	4.78 **(0.62)**	4.33 **(0.86)**	1.44 **(0.18)**	0.82 **(0.16)**
FFMI	1.43 **(0.62)**	2.45 **(0.86)**	0.40 **(0.18)**	0.48 **(0.17)**

Note: These results represent the precision of our method on completely novel input from scanning systems not represented in the training data. The precision error is worse than our results on a held-out test set of systems 1 through 3, ranging 3–5 times that of DXA.

Abbreviations: DXA, dual-energy X-ray absorptiometry; FMI, fat mass index; FFMI, fat free mass index; RMSE, root-mean-squared errors.