

RESEARCH ARTICLE

# A method for measuring human body composition using digital images

Olivia Affuso<sup>1,2\*</sup>, Ligaj Pradhan<sup>3</sup>, Chengcui Zhang<sup>3</sup>, Song Gao<sup>3</sup>, Howard W. Wiener<sup>1</sup>, Barbara Gower<sup>2,4</sup>, Steven B. Heymsfield<sup>5</sup>, David B. Allison<sup>6\*</sup>

**1** Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, United States, **2** Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, United States, **3** Department of Computer and Information Sciences, College of Arts and Sciences, University of Alabama at Birmingham, Birmingham, AL, United States, **4** Department of Nutrition Science, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States, **5** Pennington Biomedical Research Center, Louisiana State University System, Baton Rouge, LA, United States, **6** Department of Epidemiology and Biostatistics, School of Public Health, Indiana University-Bloomington, Bloomington, IN, United States

\* [Oaffuso@uab.edu](mailto:Oaffuso@uab.edu) (OA); [Allison@iu.edu](mailto:Allison@iu.edu) (DBA)



**OPEN ACCESS**

**Citation:** Affuso O, Pradhan L, Zhang C, Gao S, Wiener HW, Gower B, et al. (2018) A method for measuring human body composition using digital images. *PLoS ONE* 13(11): e0206430. <https://doi.org/10.1371/journal.pone.0206430>

**Editor:** Rebecca A. Krukowski, University of Tennessee Health Science Center, UNITED STATES

**Received:** September 12, 2017

**Accepted:** October 12, 2018

**Published:** November 5, 2018

**Copyright:** © 2018 Affuso et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Access to a de-identified dataset required to replicate our results is available via OPEN ICPSR ([www.openicpsr.org](http://www.openicpsr.org)) – Project ID: OPENICPSR-101343 (<http://doi.org/10.3886/E101343V1>).

**Funding:** This research was supported by the National Heart, Lung, and Blood Institute and the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Numbers R01HL107916 and P30DK056336. The content is solely the responsibility of the authors and does not

## Abstract

### Background/Objectives

Body mass index (BMI) is a proxy for obesity that is commonly used in spite of its limitation in estimating body fatness. Trained observers with repeated exposure to different body types can estimate body fat (BF) of individuals compared to criterion methods with reasonable accuracy. The purpose of this study was to develop and validate a computer algorithm to provide a valid estimate %BF using digital photographs.

### Subjects/Methods

Our sample included 97 children and 226 adults (age in years: 11.3±3.3; 38.1±11.6, respectively). Measured height and weight were used (BMI in kg/m<sup>2</sup>: 20.4±4.4; 28.7±6.6 for children and adults, respectively). Dual x-ray absorptiometry (DXA) was the criterion method. Body volume (BV<sub>PHOTO</sub>) and body shape (BS<sub>PHOTO</sub>) were derived from two digital images. Final support vector regression (SVR) models were trained using age, sex, race, BMI for %BF<sub>NOPHOTO</sub>, plus BV<sub>PHOTO</sub> and BS<sub>PHOTO</sub> for %BF<sub>PHOTO</sub>. Separate validation models were used to evaluate the learning algorithm in children and adults. The differences in correlations between %BF<sub>DXA</sub>, %BF<sub>NOPHOTO</sub> and %BF<sub>PHOTO</sub> were tested using the Fisher's Z-score transformation.

### Results

Mean BF<sub>DXA</sub> and BF<sub>PHOTO</sub> were 27.0%±9.2 vs. 26.7%±7.4 in children and 32.9±10.4% vs. 32.8%±9.3 in adults. SVR models produced %BF<sub>PHOTO</sub> values strongly correlated with %BF<sub>DXA</sub>. Our final model produced correlations of  $r_{DP} = 0.80$  and  $r_{DP} = 0.87$  in children and adults, respectively for %BF<sub>PHOTO</sub> vs. %BF<sub>DXA</sub>. The correlation between %BF<sub>NOPHOTO</sub> and %BF<sub>DXA</sub> was moderate, yet statistically significant in both children  $r_{DB} = 0.70$ ;  $p < 0.0001$  and adults  $r_{DB} = 0.86$ ;  $p < 0.0001$ . However, the correlations for  $r_{DP}$  were statistically higher than

necessarily represent the official views of the National Institutes of Health or any other organization.

**Competing interests:** No honoraria, grants, or other forms of payment were given to anyone to produce this manuscript. However, Dr. Allison has received grants, honoraria, donations, and consulting fees from numerous food, beverage, pharmaceutical companies, and other commercial, government, and nonprofit entities with interests in obesity. None of Dr. Allison's sponsors were involved in the design, data collection, analysis, interpretation, or decision to publish the results presented in this manuscript. The remaining authors of this manuscript have no competing interests to disclose. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

$r_{DB}$  (%BF<sub>DXA</sub> vs. %BF<sub>NOPHOTO</sub>) in both children and adults (children:  $Z = 5.95$ ,  $p < 0.001$ ; adults:  $Z = 3.27$ ,  $p < 0.0001$ ).

## Conclusions

Our photographic method produced valid estimates of BF in both children and adults. Further research is needed to create norms for subgroups by sex, race/ethnicity, and mobility status.

## Introduction

Assessment of body composition, particularly fat and fat-free mass, is vital to understanding many health-related conditions including cachexia induced by HIV, cancer, and other diseases; multiple sclerosis; wasting in neurological disorders such as Parkinson's, Alzheimer's, and muscular dystrophy; sarcopenia; obesity; eating disorders; proper growth in children, and response to exercise [1–7]. Nevertheless, challenges remain in the determination of these aspects of body composition in research [8]. Obesity, characterized by an excess of body fat (BF) and sarcopenia, defined as diminution of primarily skeletal muscle, remain significant public health problems [9, 10]. Both obesity and sarcopenia can be assessed using highly accurate techniques such as dual-energy x-ray absorptiometry (DXA) or magnetic resonance imaging (MRI) but are not widely used in large-scale epidemiologic studies or non-clinical settings due, in part, cost and size of the equipment used for these methods. Furthermore, field methods such as multiple skinfold measurements depend heavily upon repeated training of research staff to obtain accurate and reliable assessments [11]. Therefore, body mass index (BMI; kg/m<sup>2</sup>) is a commonly used alternative, but is limited in that it is an assessment of body weight relative to height and not of body composition per se. It is well-documented that obesity is often misclassified when BMI is used as a proxy for body fatness compared to measurement techniques via imaging. [12, 13]. In children, age- and sex-specific BMI percentiles have been found to underestimate the prevalence of excess adiposity when compared to DXA particularly among whites and Mexican American youth [12]. Similarly in adults, BMI misclassified obesity status differentially by race/ethnicity and age [13]. Therefore, there is significant need for a simple, portable, and relatively inexpensive but accurate measurement of body composition that performs well across age, sex, and racial/ethnic groups.

The use of digital photography may be a viable alternative to BMI for the assessment of human body composition in field research. This method has the potential to overcome limitations associated with BMI particularly the misclassification of obesity among individuals who have relatively high lean mass (i.e. body builders) or low lean mass (i.e. the elderly). Our approach to using digital images to assess human body composition is predicated on evidence from studies done as early as the 1930's using either visual estimation [11, 14–16]; or photographic assessment of body volume, from which body composition could be determined [17–19]. Previous studies employing visual estimation have found that both trained and untrained observers can provide moderately accurate estimates of percentage BF by visually inspecting an individual directly or from photography with correlations between observations and criterion measures, such as under water weighing (UWW) ranging between  $r = 0.56$  and  $0.83$ . This evidence suggests that visual estimation of body composition can be valid, but may be limited by familiarity of the observer with the study population, subjectivity of the observer, and reproducibility of the study results. Prior to the advent of digital photography, researchers used

manual photographic methods to provide valid and reliable estimates of body volume as a means to overcome the substantial expense and participant burden associated with UWW. However, these early attempts still required significant labor to process the photographs and manually calculate body volume. The automation of the visualization process by using computerized digital image analysis could overcome several of the issues associated with visual assessment of body volume and composition. Therefore, the purpose of this study was to develop an easy, portable, quick, and comparatively inexpensive, but valid computerized image analysis method for use in large-scale and/or remote studies to estimate fat and fat free mass.

## Methods

### Study sample

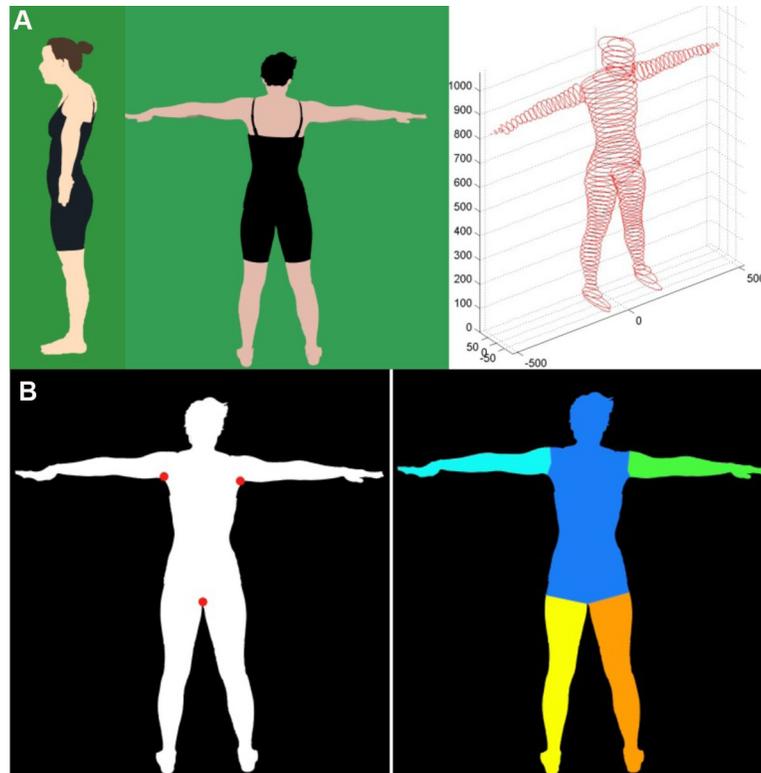
Participants were 323 children and adults aged 6–80 years representing a broad range of shapes and sizes recruited from the metropolitan Birmingham, AL area (2012–2014) via flyers, newspapers, newsletters, an online research referral service, word of mouth, visits to local recreational centers, churches, and community events. Inclusion criteria were: 1) without diseases known to effect body composition; 2) able to stand for three photographs; 3) no contraindication for a DXA body composition scan; 4) not missing more than one finger or toe to reduce error in volume determination; and 5) not pregnant. Participants were instructed to refrain from caffeine and large meals prior to the single study visit. Informed consent was obtained from the adults/parents and children provided their assent for study participation. The study protocol was approved by the Institutional Review Board of The University of Alabama at Birmingham. All participants received \$20 for completing the study.

### Measures

All measures were assessed by trained research staff with participants dressed in close fitting but non-compressing LYCRA shorts and tank/sports bra (females only) without shoes. Height (to the nearest 0.1 cm) and weight (to the nearest 0.1 kg) were measured using a physician's balance beam scale with stadiometer (HealthOMeter—Model 402LB, McCook, IL). BMI ( $\text{kg}/\text{m}^2$ ) was calculated from measured height and weight. Obesity status was categorized in children as BMI  $\geq 95$  percentile and BMI  $\geq 30$  in adults, according to standard definition by expert panels [20, 21]. Body composition was measured using DXA (GE Lunar iDXA, Madison, WI) with the pediatric software employed when appropriate (encore 2011 version 13.6). Quality assurance tests were performed daily per the manufacturer instructions. Obesity via DXA was defined as body fat percent  $\geq 25\%$  in males,  $\geq 30\%$  girls and  $\geq 35\%$  in women [22, 23]. Three photographs (front, back, and side profiles) were taken with a digital camera (Canon PowerShot—Model SX50; Canon USA Inc., Melville, NY) with participants standing against a photography green screen.

### Body volume and body shape from two-dimensional image processing

Body volume ( $BV_{\text{PHOTO}}$ ) and body shape ( $BS_{\text{PHOTO}}$ ) were estimated from photographs of each participant using two photographs (back and side profiles). Although we had a third photograph of the front profile, the use of the additional photograph did not change our estimation of body volume or shape. Briefly, the methods used to determine BV and BS are described below. However, our technical report of the detailed methods used herein are described elsewhere [24].



**Fig 1.** A. 3D-body model from 2D images. B. Key points to separate body segments.

<https://doi.org/10.1371/journal.pone.0206430.g001>

We constructed a 3-dimensional body model based on both the back and side body masks (i.e., outline of the body) extracted from 2D profile images as shown in Fig 1A. The body position for the back profile photograph required participants to stand with their arms and legs separated, while for the side profile photograph the arms were required to be close to the body with the legs together. Also for the side profile, the right leg and foot were covered in a green cloth to isolate the side of the body. The distance of the camera from participants (91 in.) and the camera/lighting settings were standardized to reduce variation in the photographs among participants. A four-step procedure was then used to separate the body components of each participant and was as follows: 1) The body was separated from the green screen by setting a color intensity threshold that facilitated the differentiation between the background and the black clothing and skin tone in the photograph; 2) The body mask from the back and side profiles were rotated until symmetrical in the vertical plane; 3) The height from the body mask was then normalized to represent participants' actual height; 4) Finally, the separation points, called *key points* (see Fig 1B), used to determine the separation line of each body component such as the arms, legs and the trunk, were detected for each participant from the back profile, which allowed better separation due to natural folds of the skin.

The local dimension features (e.g. length and width) of each body component such as arms, legs, and trunk, were used to construct ellipse-like slices along the main orientation of each component. Long and short axes were calculated for each ellipse-like slice depending upon the length and width of each body component. The area of each slice was taken to be the number of pixels within that slice. A 3D body model was then constructed by accumulating ellipse-like slices and the  $BV_{\text{PHOTO}}$  was derived by summing the area of all slices.

Body shape features, which delineate fat distribution in the trunk, were captured by extracting the front curve (FC) and the side curve (SC) using the left upper key point and the lower key point as demonstrated in Fig 2. The width of the body was taken at each of 12 equally spaced points along the vertical contour of the front of the body, providing 12 numerical values representing the FC of each person. Likewise, the side curve of the body was measured by 12 additional equally spaced lines representing the extracted side contour. The side contour lines were measured to the central bodyline (CBL, formed by a vertical line drawn from the top of the head through the lower key point) and gave the twelve numerical values representing each person's SC. Details of the methods used to extract the shape features are published elsewhere [24].

After extracting FC and SC, a cluster evaluation function in MatLab ('evalclusters') with K-means clustering and 'Calinski-Harabasz' criterion was used to compute optimal number of clusters 'k' for both the FC and SC features from the training dataset [25]. This cluster evaluation step examined cluster sizes from 2 to 10 and computed the optimal cluster size for the FC and SC feature sets. FC and SC were then represented as *k*-element vectors, consisting of 0's and 1's. A value of one at the *n*<sup>th</sup> element indicates that the body shape is in the *n*<sup>th</sup> cluster. The encoded vectors for FC and BC are represented as  $BS_{\text{PHOTO}}$  in our model and used as categorical features to train the prediction models. After determining the optimal number of clusters and their corresponding cluster-centroids using only the training dataset, FC and SC in the testing dataset were later assigned to the clusters whose centroids were closest in terms of Euclidean distance.

## Building the prediction model

$BV_{\text{PHOTO}}$  and  $BS_{\text{PHOTO}}$  of each participant along with the covariates age, race, sex, BMI were used as input features as shown in the first step of Table 1, to train a support vector regression (SVR) for the estimation of  $\%BF_{\text{PHOTO}}$ . The ground truth collection, type of regression model used and the parameter optimization for the regression model are also described in steps two, three and four of Table 1. Separate 3-fold cross validation models were used to evaluate the performance of the learning algorithm for the prediction of body fat in children and adults [26,27]. Both datasets were then shuffled randomly and divided into three equal subsets. The 3-fold cross validations were then conducted by taking two of the subsets as the training dataset and remaining one subset as the testing dataset without any overlap. In addition, we repeated the process three times, choosing separate subsets for testing each time. Hence, we

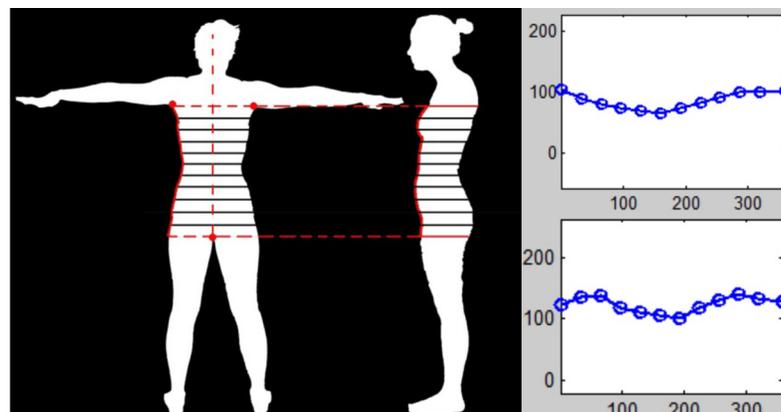


Fig 2. Body shape feature derives from front and side curves.

<https://doi.org/10.1371/journal.pone.0206430.g002>

**Table 1. Procedures used in the support vector regression for the photographic estimation of body fat.**

Steps	Detail description						
<b>1. Feature set generation</b>	<p>Back and side profile images of 323 children and adults ranging from age 6–80 were collected along with their age, sex, race and BMI information. The pictures are used to compute <math>BV_{PHOTO}</math>, FC and SC. FC and SC were separately used to cluster the participants into <math>k</math> groups which is computed using a cluster evaluation function in MatLab ('evalclusters') with K-means clustering and 'Calinski-Harabasz' criterion. These clusters represent different body shapes and are represented as <math>BS_{PHOTO}</math> in our study. Hence, our feature set for each participant is:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>age</td> <td>sex</td> <td>race</td> <td>BMI</td> <td><math>BV_{PHOTO}</math></td> <td><math>BS_{PHOTO}</math></td> </tr> </table>	age	sex	race	BMI	$BV_{PHOTO}$	$BS_{PHOTO}$
age	sex	race	BMI	$BV_{PHOTO}$	$BS_{PHOTO}$		
<b>2. Ground truth</b>	We also measure the BF% for each participant using a DXA machine ( $BF_{DXA}\%$ ). $BF_{DXA}\%$ will be our ground truth BF% for training and testing the prediction model.						
<b>3. Regression model</b>	We train Support Vector Regression (SVR) model to predict $BF_{DXA}$ using our feature set. We use the 'nu-SVR' type SVR with 'radial bias' type kernel function from LIBSVM[31]						
<b>4. Parameter optimization</b>	'nu-SVR' requires several parameters [31] Values of -s and -t are selected to be 4 and 2 respectively. A parameter selection during operation was performed using our training dataset to select the best parameter set for the -g (gamma), -c (cost) and -n (nu). Combinations of g, c and n, within certain ranges were tested training, and the best combination that produced the highest prediction accuracy in the training data set was chosen for our regression model. The values of -g and -c were taken as $2^X$ , where X is an integer ranging from -10 to 10. Similarly values of n were taken from 0.1 to 1 at an interval of 0.1.						
<b>5. Training and Testing</b>	The participants were randomly divided into 3 subsets and 3-fold cross validations were performed by taking two of the subsets as the training dataset and remaining subset as the testing dataset. FC and SC curves from the training dataset were used to compute the $k$ clusters. Each test dataset was assigned to the clusters discovered during training by computing the nearest cluster centroid to compute their $BS_{PHOTO}$ . This process was repeated 3 times, choosing separate subsets for testing each time.						

<https://doi.org/10.1371/journal.pone.0206430.t001>

tested each data point in the dataset once after completing one round of the 3-fold cross validation process as described in step 5 of Table 1.

### Statistical analysis

Descriptive statistics (means, SD, ranges, and/or frequencies) were computed for the participant characteristics stratified by adults and children.

Pearson's correlations between the DXA criterion measure of body fat and predicted values from a simplified model ( $BV_{PHOTO}$  + covariates) and the full ( $BV_{PHOTO}$  +  $BS_{PHOTO}$  + covariates) model were calculated. The correlations are represented as follows:

Where,

$r_{DP}$  denotes the correlation between %BF by DXA and % BF by PHOTO,

$r_{DB}$  denotes the correlation between %BF by DXA and %BF by BMI plus covariates.

We used the method of Meng et al.[28] to determine whether our method (i.e. % $BF_{PHOTO}$ ) was significantly better correlated to the criterion method of % $BF_{DXA}$  than was % $BF_{NO-PHOTO}$ [29]. We also created Bland-Altman plots of the difference in estimation of body fat by each method to examine the error in our estimation across the range of % $BF_{DXA}$  and %  $BF_{PHOTO}$  measurements [30].

SVR models were constructed using LIBSVM[31] in MatLab Version R2014b (Mathworks, Inc., Natick, MA). Correlation and the test for statistical difference were computed using SAS Version 9.3 (SAS, Inc., Cary, NC). Alpha level of significance was set at  $p < 0.05$ , 2-tailed).

### Results

Participant characteristics stratified by children (6–18 years) and adults ( $\geq 19$  years) are presented in Table 2. Our sample included 97 children and 226 adults (mean  $\pm$  SD of age in years:  $11.3 \pm 3.3$ ;  $38.1 \pm 11.6$ , respectively). Mean BMI and % $BF_{DXA}$  were  $20.4 \text{ kg/m}^2 \pm 4.4$ ;  $28.7 \text{ kg/m}^2 \pm 6.6$  and  $27.1\% \pm 9.2$ ;  $32.7\% \pm 10.4$  for children and adults, respectively). Among the children,

**Table 2. Sample characteristics.**

	Adults n = 226	Children n = 97
Age (mean, SD), years	38.1 (11.1)	11.3 (3.3)
Percentage Female (n, %)	105 (46.5)	46 (47.4)
Percentage African American (n, %)	105 (46.5)	59 (60.8)
Height (mean, SD), cm	169.2 (8.4)	148.5 (15.4)
Weight (mean, SD), kg	82.2 (20.1)	46.3 (16.2)
BMI (mean, SD), kg/m <sup>2</sup>	28.7 (6.6)	20.4 (4.4)
BMI Percentile (mean, SD)	-	64.8 (27.6)
DXA Body Fat (mean, SD), %	32.9 (10.4)	27.0 (9.2)
Photographic Volume (MP)	25.4 (6.3)	14.1 (4.9)

MP = megapixel

<https://doi.org/10.1371/journal.pone.0206430.t002>

47.4% were female, 60.8% were African American (AA), while among the adults, 46.5% were female and 46.5% were AA. Also, 14.4% of children and 33.6% of adults were classified as obese using BMI (age- and sex-specific BMI  $\geq 95^{\text{th}}$  percentile for children and  $\geq 30 \text{ kg/m}^2$  for adults), while 40.2% of children and 66.4% of adults were over the threshold for body fatness via DXA ( $\geq 30\%$  for girls and  $\geq 25\%$  for boys;  $\geq 35\%$  for women and  $\geq 25\%$  for men).

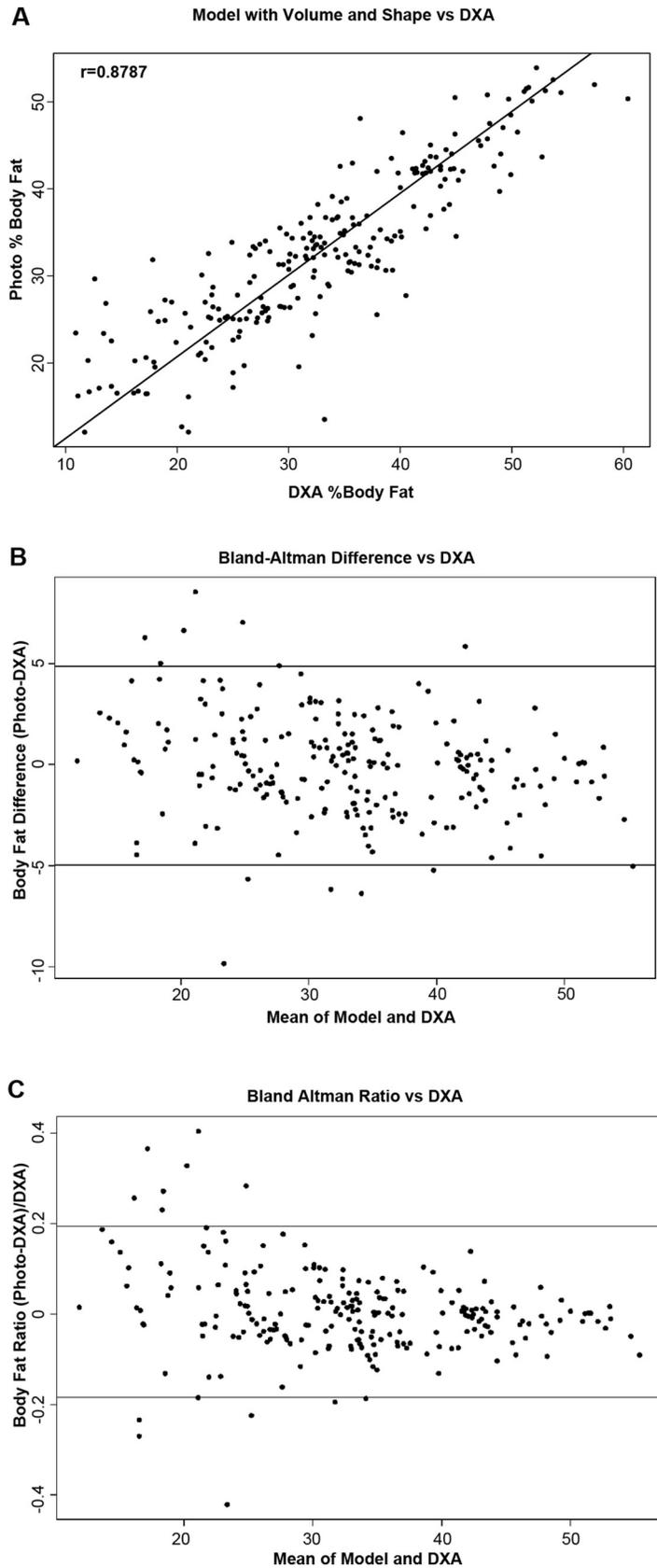
Correlations between %BF<sub>DXA</sub> and the %BF<sub>PHOTO</sub> from our learning algorithms are presented for adults and children in Figs 3 and 4, respectively. Both SVM models produced %BF<sub>PHOTO</sub> values that were strongly correlated with the %BF<sub>DXA</sub>. %BF<sub>PHOTO</sub> from Model 1 which included demographic variables plus BMI and BV<sub>PHOTO</sub> produced correlations of  $r_{DP} = 0.72$  and  $r_{DP} = 0.86$  in children and adults respectively. However, %BF<sub>PHOTO</sub> from Model 2 which included variables from Model 1 plus BS<sub>PHOTO</sub> produced  $r_{DP} = 0.81$  in children and  $r_{DP} = 0.88$  in adults.

The correlations among %BF<sub>DXA</sub>, %BF<sub>NOPHOTO</sub>, and %BF<sub>PHOTO</sub> are reported in Table 3. The correlation between %BF<sub>NOPHOTO</sub> and %BF<sub>DXA</sub> was moderate, yet statistically significant in both children ( $r_{DB} = 0.70$ ;  $p < 0.0001$ ) and adults ( $r_{DB} = 0.86$ ;  $p < 0.0001$ ).  $r_{DP}$  was significantly larger than  $r_{DB}$  in both children and adults (children:  $Z = 5.95$ ;  $p < 0.0001$ ; adults:  $Z = 3.27$ ;  $p < 0.001$ ). Bland-Altman plots revealed greater differences between our %BF<sub>PHOTO</sub> and the %BF<sub>DXA</sub> in the tails of the distributions in both adults and children (Fig 3A and 3B). In adults, the mean absolute and relative differences between the two methods were -0.06 (95% CI: -4.97, 4.85) and 0.005 (95% CI: -0.18, 0.19), respectively. In children, the mean and absolute and relative differences were -0.19 (95% CI: -5.43, 5.04) and 0.004 (95% CI: -0.20, 0.21), respectively.

## Discussion

Innovations in both digital photography and image analysis have allowed the automation of visual body composition assessment. Our study results showed strong correlations between predicted body fat percentage derived from 2D digital photographs and body fatness from DXA in a sample of children and adults. The performance of our image analysis algorithm provided statistically better correlations to DXA measurements than containing BMI and demographic information. The average absolute error between our method and DXA was small (~4.1%) with BV and BS contributing significantly to the overall prediction of percent body fat.

Our estimates of body volume from the photographs were highly correlated with body volume from air displacement plethysmography suggesting that our photographic method is



**Fig 3. Correlations between predicted body fat from the photographs and dual-energy x-ray absorptiometry (DXA).** (A) and Bland-Altman plots of the absolute (B) and relative (C) differences between the two methods (horizontal lines represent the 95% confidence intervals) among adults.

<https://doi.org/10.1371/journal.pone.0206430.g003>

valid for use in the prediction of body composition ( $r = 0.98$ ) [32]. To further enhance our estimation of percent body fat, we incorporated visual elements of body shape that were also shown to correspond to levels of adiposity from DXA (i.e. thin body shapes corresponded to low body fatness). The clustering of body shape and percent body fat suggests that body shape provides important visual information for improving the accuracy of the image analysis algorithm and may have implications on the assessment of health outcomes beyond overall adiposity. The findings followed a similar pattern among children and adults though the correlation between our method and DXA was higher in adults.

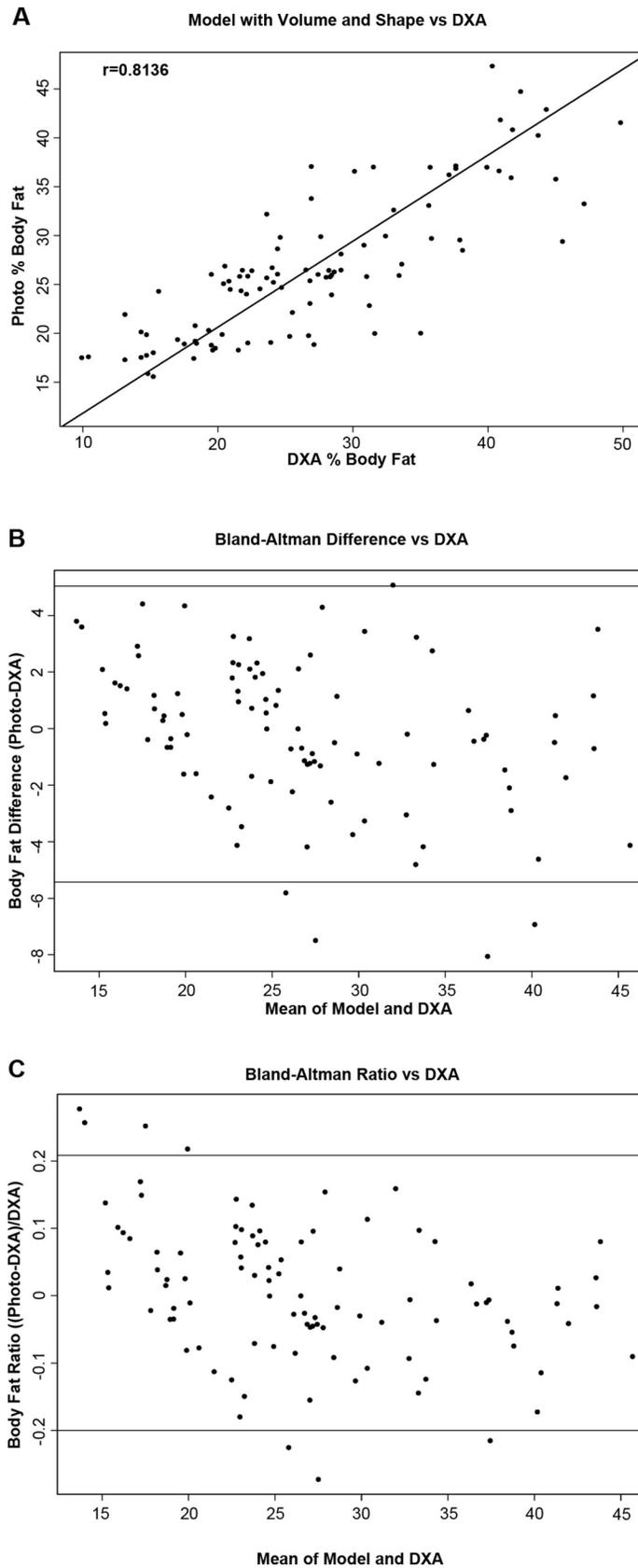
Moreover, our results are consistent with other studies that use multiple cameras (e.g. 8–16 cameras) to estimate BF in humans [33, 34]. The simplicity of our method overcomes several issues associated with other field measures of body composition such as portability, ease of data collection, and expense. Relative to a DXA machine (~\$30,000 USD), the setup cost for our photographic method is ~\$300, which includes the cost for a scale/stadiometer, Chroma-key green background, tank top & shorts, and digital camera. Our preliminary software program has been developed in Matlab and will require the end-user to enter demographic information along with the selection of a front and side photograph previously saved on the computer. Furthermore, our method requires little human input during the image capture and processing procedures, thereby reducing biases inherent in other field methods such as visual estimation and skinfold measurement, which require training and re-training [11, 14–16, 35]. Lastly, no specialized equipment is required beyond a simple digital camera and a green backdrop makes our portable method more convenient than other methods.

## Implications

Computerized image analysis of digital photographs can be used as a valid method for estimating body fatness in humans. Simple digital photographs processed with our algorithm could be used in place of BMI alone in both clinical practice and public health research. This is of particular interest in research as evidence suggests that associations between obesity assessed by BMI and outcomes such as mortality may be different (e.g. Linear vs U-shaped) when a more robust method of body composition is used. Our novel method also addresses the need for a portable, yet valid, method for measuring body composition in large epidemiological studies as well as investigations conducted in remote locations.

## Strengths & limitations

The strengths of this study include the use of digital photographs to capture unbiased information about the body for use in the estimation of volume and shape, DXA as the criterion measure of body composition, and a diverse sample of black and white youth and adult males and females. However, several limitations of this work should be also noted, including the small sample within each stratum of race, sex and age as well as the inclusion of only able-bodied individuals who could stand for the photographs. Although attempts were made to include a broad range of body sizes at this stage in the development of this method, there were limited numbers of participants enrolled in the tails of the distribution for this analysis. Therefore, the results are not necessarily generalizable to very lean or very obese individuals, all race-ethnic groups or persons unable to stand for the photographs. Future studies should examine the



**Fig 4. Correlations between predicted body fat from the photographs and dual-energy x-ray absorptiometry (DXA).** (A) and Bland-Altman plots of the absolute (B) and relative (C) differences between the two methods (horizontal lines represent the 95% confidence intervals) among children.

<https://doi.org/10.1371/journal.pone.0206430.g004>

**Table 3. Correlations and concordance between DXA and estimated body fat.**

	Adults n = 226		Children n = 97	
	DXA % Body Fat	Lin's Concordance Coefficient	DXA % Body Fat	Lin's Concordance Coefficient
Non-Photographic % Body Fat	0.86 <sup>a</sup>	0.85	0.70 <sup>HYPERLINK a</sup>	0.66
Photographic % Body Fat (Volume)	0.86	0.85	0.72	0.67
Photographic % Body Fat (Volume and Shape)	0.88 <sup>b</sup>	0.87	0.81 <sup>b</sup>	0.79

DXA = dual energy x-ray absorptiometry

<sup>a,b</sup>Differences in Correlation Coefficients between non-photographic model and photographic model containing volume and shape: Adults—Z = 3.27, P<0.001; Children—Z = 5.95, P<0.0001. Lin's Concordance Coefficients are comparisons of the model estimates of % body fat and DXA % body fat.

<https://doi.org/10.1371/journal.pone.0206430.t003>

performance of our photographic method in a larger sample to allow stratification by age, sex, and race/ethnicity.

## Conclusion

This research shows that a computer algorithm can be developed to provide a valid estimate of body fatness from 2D photographs taken with a regular digital camera.

## Acknowledgments

We wish to thank our study participants, without whom this research would not have been possible.

## Author Contributions

**Conceptualization:** Olivia Affuso, Steven B. Heymsfield, David B. Allison.

**Data curation:** Olivia Affuso, Ligaj Pradhan, Chengcui Zhang, Song Gao, Howard W. Wiener.

**Formal analysis:** Olivia Affuso, Howard W. Wiener, David B. Allison.

**Funding acquisition:** Olivia Affuso, David B. Allison.

**Investigation:** Olivia Affuso.

**Methodology:** Olivia Affuso, Ligaj Pradhan, Chengcui Zhang, Song Gao, Barbara Gower, Steven B. Heymsfield, David B. Allison.

**Project administration:** Olivia Affuso.

**Resources:** Barbara Gower.

**Supervision:** Olivia Affuso, Chengcui Zhang.

**Writing – original draft:** Olivia Affuso, Ligaj Pradhan, Howard W. Wiener.

**Writing – review & editing:** Olivia Affuso, Barbara Gower, Steven B. Heymsfield, David B. Allison.

## References

1. Vaisman N, Corey M, Rossi MF, Goldberg E, Pencharz P. Changes in body composition during refeeding of patients with anorexia nervosa. *J Pediatr*. 1988; 113(5):925–9. PMID: [3183854](#).
2. Johnson DK, Wilkins CH, Morris JC. Accelerated weight loss may precede diagnosis in Alzheimer disease. *Arch Neurol*. 2006; 63(9):1312–7. <https://doi.org/10.1001/archneur.63.9.1312> PMID: [16966511](#).
3. Stewart R, Masaki K, Xue QL, Peila R, Petrovitch H, White LR, et al. A 32-year prospective study of change in body weight and incident dementia: the Honolulu-Asia Aging Study. *Arch Neurol*. 2005; 62(1):55–60. <https://doi.org/10.1001/archneur.62.1.55> PMID: [15642850](#).
4. Morley JE, Thomas DR, Wilson MM. Cachexia: pathophysiology and clinical relevance. *Am J Clin Nutr*. 2006; 83(4):735–43. <https://doi.org/10.1093/ajcn/83.4.735> PMID: [16600922](#).
5. Hunter GR, McCarthy JP, Bamman MM. Effects of resistance training on older adults. *Sports Med*. 2004; 34(5):329–48. <https://doi.org/10.2165/00007256-200434050-00005> PMID: [15107011](#).
6. Allison DB, Mentore JL, Heo M, Chandler LP, Cappelleri JC, Infante MC, et al. Antipsychotic-induced weight gain: a comprehensive research synthesis. *Am J Psychiatry*. 1999; 156(11):1686–96. <https://doi.org/10.1176/ajp.156.11.1686> PMID: [10553730](#).
7. Knox TA, Zafonte-Sanders M, Fields-Gardner C, Moen K, Johansen D, Paton N. Assessment of nutritional status, body composition, and human immunodeficiency virus-associated morphologic changes. *Clin Infect Dis*. 2003; 36(Suppl 2):S63–8. <https://doi.org/10.1086/367560> PMID: [12652373](#).
8. Segal KR, Dunaif A, Gutin B, Albu J, Nyman A, Pi-Sunyer FX. Body composition, not body weight, is related to cardiovascular disease risk factors and sex hormone levels in men. *J Clin Invest*. 1987; 80(4):1050–5. <https://doi.org/10.1172/JCI113159> PMID: [3654969](#); PubMed Central PMCID: [PMCPMC442345](#).
9. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity in the United States, 2009–2010. *NCHS Data Brief*. 2012;( 82):1–8. PMID: [22617494](#).
10. Fielding RA, Vellas B, Evans WJ, Bhasin S, Morley JE, Newman AB, et al. Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. *J Am Med Dir Assoc*. 2011; 12(4):249–56. <https://doi.org/10.1016/j.jamda.2011.01.003> PMID: [21527165](#); PubMed Central PMCID: [PMCPMC3377163](#).
11. Vogel JA, Kirkpatrick JW, Fitzgerald PI, Hodgdon JA, Harman EA. Derivation of anthropometry based body fat equations for the Army's weight control program (No. USARIEM-T-17/88). Army Research Institute of Environmental Medicine, Natick, MA. 1988.
12. Affuso F, Bray MS, Fernandez JR, Casazza K. Standard obesity cut points based on BMI percentiles do not equally correspond to body fat percentage across racial/ethnic groups in a nationally representative sample of children and adolescents. *Intl J Body Compos Res*. 2010; 8(4):117–22.
13. Baumgartner RN, Heymsfield SB, Roche AF. Human body composition and the epidemiology of chronic disease. *Obes Res*. 1995; 3(1):73–95. PMID: [7712363](#).
14. Blanchard JM, Ward GM, Kryzywicki HJ, Canham JE. A visual appraisal method for estimating body composition in humans. Presidio of San Francisco, Report No.81: 1979.
15. Eckerson JM, Housh TJ, Johnson GO. The validity of visual estimations of percent body fat in lean males. *Med Sci Sports Exerc*. 1992; 24(5):615–8. PMID: [1569858](#).
16. Sterner TG, Burke EJ. Body fat assessment: A comparison of visual estimation and skinfold techniques. *Physician Sportsmed*. 1986; 14:101–7.
17. Weinbach AP. Contour maps, center of gravity, moment of inertia, and surface area of the human body. *Human Biol*. 1938; 10:356–71.
18. Pierson WR. Monophotogrammetric determination of body volume. *Ergonomics*. 1961; 4(3):213–8.
19. Geoghegan B. The determination of body measurements, surface area and body volume by photography. *Am J Phys Anthropol*. 1953; 11(1):97–119. PMID: [13040507](#).
20. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000; 320(7244):1240–3. PMID: [10797032](#); PubMed Central PMCID: [PMCPMC27365](#).
21. NHLBI Expert Panel: Obesity Education Initiative. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. 1998.
22. Williams DP, Going SB, Lohman TG, Harsha DW, Srinivasan SR, Webber LS, et al. Body fatness and risk for elevated blood pressure, total cholesterol, and serum lipoprotein ratios in children and adolescents. *Am J Public Health*. 1992; 82(3):358–63. PMID: [1536350](#); PubMed Central PMCID: [PMCPMC1694353](#).
23. Gallagher D, Heymsfield SB, Heo M, Jebb SA, Murgatroyd PR, Sakamoto Y. Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index. *Am J Clin Nutr*. 2000; 72(3):694–701. <https://doi.org/10.1093/ajcn/72.3.694> PMID: [10966886](#).

24. Pradhan L, Song G, Zhang C, Gower BA, Heymsfield SB, Allison DB, et al., editors. Feature extraction from 2D images for body composition analysis. *IEEE International Symposium on Multimedia*; 2015:45–52. Miami, FL: IEEE. <https://doi.org/10.1109/ISM.2015.117> URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7442294&isnumber=7442255>
25. Maulik U, Bandyopadhyay S. Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(12):1650–4.
26. Ivanescu AE, Li P, George B, Brown AW, Keith SW, Raju D, et al. The importance of prediction model validation and assessment in obesity and nutrition research. *Int J Obes (Lond)*. 2015. <https://doi.org/10.1038/ijo.2015.214> PMID: 26449421.
27. Ambrose C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci U S A*. 2002; 99(10):6562–6. <https://doi.org/10.1073/pnas.102102699> PMID: 11983868; PubMed Central PMCID: PMC124442.
28. Meng XL, Rosenthal R, Rubin DR. Comparing correlated correlation coefficients. *Psych Bulletin*. 1992; 111(1):172–5.
29. Lin L, Hedayat AS, Wu W. *Statistical tools for measuring agreement* Springer Science and Business Media; 2012.
30. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet*. 1995; 346(8982):1085–7. PMID: 7564793.
31. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol*. 2011; 2(3):1–27. <https://doi.org/10.1145/1961189.1961199>
32. Affuso O, Zhang C, Chen W-B, Song G, Keeting K, Lewis DW, et al. Novel image analysis method for assessing body composition in humans: a pilot study *Proceedings of the 31st Annual Scientific Meeting of the Obesity Society*, Atlanta, GA. 2013.
33. Connell L, Ulrich P, Brannon E, Alexander M, Presley A. Body shape assessment scale: Instrument development for analyzing female figures. *Clothing and Textile Res J*. 2006; 24:80–95.
34. Xu B, Yu W, Yao M, Pepper MR, Freeland-Graves JH. Three-dimensional surface imaging system for assessing human obesity. *Opt Eng*. 2009; 48(10):nihpa156427. <https://doi.org/10.1117/1.3250191> PMID: 19966948; PubMed Central PMCID: PMC12788969.
35. Pollock ML, Jackson AS. Research progress in validation of clinical methods of assessing body composition. *Med Sci Sports Exerc*. 1984; 16(6):606–15. PMID: 6392815.