Clinical Nutrition 41 (2022) 144–152

Contents lists available at ScienceDirect

Clinical Nutrition

journal homepage: http://www.elsevier.com/locate/clnu

Original article

Artificial neural network model effectively estimates muscle and fat mass using simple demographic and anthropometric measures

Prabhat Pathak [a], Siddhartha Bikram Panday [b], Jooeun Ahn [a, c, *]

[a] Department of Physical Education, Seoul National University, Republic of Korea
[b] Department of Sports and Leisure Studies, Keimyung University, Republic of Korea
[c] Institute of Sport Science, Seoul National University, Republic of Korea

ARTICLE INFO

Article history:
Received 11 August 2021
Accepted 21 November 2021

Keywords:
Lean muscle mass
Fat mass
Artificial neural network
Multiple linear regression

SUMMARY

Background & aims: Lean muscle and fat mass in the human body are important indicators of the risk of cardiovascular and metabolic diseases. Techniques such as dual-energy X-ray absorptiometry (DXA) accurately measure body composition, but they are costly and not easily accessible. Multiple linear regression (MLR) models have been developed to estimate body composition using simple demographic and anthropometric measures instead of expensive techniques, but MLR models do not explore nonlinear interactions between inputs. In this study, we developed simple demographic and anthropometric measure-driven artificial neural network (ANN) models that can estimate lean muscle and fat mass more effectively than MLR models.
Methods: We extracted the demographic, anthropometric, and body composition measures of 20,137 participants from the National Health and Nutrition Examination Survey conducted between 1999 and 2006. We included 13 demographic and anthropometric measures as inputs for the ANN models and divided the dataset into training and validation sets (70:30 ratio) to build and cross-validate the models that estimate lean muscle and fat mass, which were originally measured using DXA. This process was repeated 100 times by randomly dividing the training and validation sets to eliminate any effect of data division on model performance. We built additional models separately for each sex and ethnicity, older individuals, and people with underlying diseases. The coefficient of determination (R^2) and standard error of estimate (*SEE*) were used to quantify the goodness of fit.
Results: The ANN models yielded high R^2 values between 0.923 and 0.981. These values were significantly higher than those of the MLR models ($p < 0.001$) in all cases. The percentage difference in R^2 between the ANN and MLR models ranged between 0.40% ± 0.02% and 2.65% ± 0.27%. The *SEE* values of the ANN models, which were below 2 kg for all cases, were significantly lower than those of MLR models ($p < 0.001$). The percentage difference in *SEE* values between the ANN and MLR models ranged between −5.67% ± 0.39% and −22.32% ± 1.98%.
Conclusions: We developed and validated an inexpensive but effective method for estimating body composition using easily obtainable demographic and anthropometric data.

1. Introduction

Lean muscle and fat are major components of the human body [1–3], and the mass of each serves as an important indicator of physical function. In general, individuals with low muscle and high fat mass are considered to be at an increased risk of cardiovascular and metabolic diseases, falls, and mortality [4–6]. Hence, it is crucial to accurately measure the amount of lean muscle and fat mass in the human body. The two most accurate human body composition measurement techniques are dual-energy X-ray absorptiometry (DXA) and magnetic resonance imaging (MRI) [7]. However, these techniques are expensive in terms of cost, space, and time. The resulting limited accessibility of these technologies has motivated researchers to explore less expensive but effective methods to estimate lean muscle and fat mass.

* Corresponding author. Seoul National University, Address: Room 412, Building 71-1, 1 Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea. Fax: +82 2 872 2867.
E-mail address: ahnjooeun@snu.ac.kr (J. Ahn).

https://doi.org/10.1016/j.clnu.2021.11.027

A large body of literature has proposed simple demographic and anthropometric measure-driven multiple linear regression (MLR) models based on DXA or MRI body composition reference data to estimate lean muscle and fat mass. In these studies, age, sex, and ethnicity were generally used as demographic measures, whereas height, weight, limb circumference, limb length, and skin thickness were used as anthropometric measures. Ross et al. developed one of the earliest demographic and anthropometric measure-driven MLR models to estimate lean muscle and fat mass, with a high goodness of fit [8]. Other studies have developed similar models to estimate lean muscle and fat mass for specific age groups [9–11], groups with different body mass indices [12], mono-ethnic populations [13–15], and elite athletes [16] with a moderate to high goodness of fit. However, the samples used to build these models were either small (<1500 individuals) or limited to a homogeneous population, diminishing the generalizability of the model.

To the best of our knowledge, Lee et al. developed and validated MLR models to estimate lean muscle and fat mass with a high goodness of fit using a (largest) sample size of 14,065 individuals [17]. The body composition data used to develop the model were recorded using DXA and collected through the National Health and Nutrition Examination Survey (NHANES) [18]. The survey collected a large population dataset with diverse demographics and a wide anthropometric range, making the model developed by Lee et al. the most generalizable to date. Nevertheless, the generalizability and efficacy of the model can be improved even further if we can detect any nonlinear interaction between input variables and iterate the model, to minimize the error between the estimated and actual output values. An artificial neural network (ANN) model is an effective option for actualizing this improvement [19–21]. ANN models mimic the learning dynamics of the human brain by formulating hidden layers that detect nonlinear interactions between input variables. Additionally, the back-propagation property of the ANN model enables robust estimation of the output values by reducing noisy input patterns and errors. Multiple studies have shown that ANN models perform better than various regression models in estimating physiological parameters [22,23] and clinical outcomes [24,25]. However, to the best of our knowledge, no prior study has used an ANN model to estimate lean muscle and fat mass using demographics and anthropometry datasets of a large population.

In this study, we analyzed the NHANES dataset and developed a demographic and anthropometric measure-driven ANN model to estimate lean muscle and fat mass. We found that the goodness of fit of the ANN model was significantly higher than that of the previous MLR model based on the same measures. This finding shows that the developed ANN model is an enhanced, alternative method for estimating lean muscle and fat mass when expensive equipment is not available.

2. Materials and methods

2.1. Data source and study population

We extracted the demographic, anthropometric, and body composition measures of participants from the NHANES conducted between 1999 and 2006 by the Center for Disease Control and Prevention and National Center of Health Statistics (NCHS) [18]. All participants were non-institutionalized citizens of the United States of America (USA). We did not merge the datasets in such a manner that the participants could be identified. Sex, age, and ethnicity were included as demographic variables. We excluded the data of participants older than 84 years because such individuals were uniformly coded as 85; including this group without information on the actual age could reduce model performance.

We included height; weight; length of the upper arm and leg; circumference of the arm, waist, thigh, and calf; and the triceps and subscapular skinfolds as anthropometric measures. Participants with self-reported weight and height above 300 pounds and 6 feet 5 inches, respectively, were excluded because the measurement table of DXA could not fit such participants. We also excluded participants with incomplete data.

Additionally, we included lean muscle and fat mass recorded using DXA as body composition measures. Five imputations were performed for each participant. Complying with the guidelines provided by the NHANES [26], we analyzed each imputation and averaged the outcomes separately instead of combining the raw data of the five imputations. Hence, we removed records of any participant with missing body composition data in any of the imputations. Finally, we selected the data of 20,137 participants (11,319 men and 8818 women) for further analysis.

2.2. Data collection

2.2.1. Demographic measures

Trained interviewers collected all demographic measures via household interviews during the period of data collection. Sex was coded as 1 for males and 2 for females. Their ages ranged from 8 to 84 years, and ethnicity was divided into five categories: Mexican American, other Hispanic, white, black, and other ethnicities (including multi-ethnic individuals), coded as 1 to 5, respectively.

2.2.2. Anthropometric measures

Following the guidelines provided by Loman et al. [27], trained health technicians measured all anthropometric measures inside a mobile examination center. A stadiometer was used to measure participant height to the nearest 0.1 cm, and a Toledo digital scale was used to measure their weights to the nearest 0.1 kg. The limb length and circumference were measured using a measuring tape to the nearest 0.1 cm. Arm and leg measurements were performed on the right side of the body.

Upper arm length was measured as the length from the acromion process to the tip of the olecranon process while asking participants to stand up, keeping their back straight, and maintaining a 90° elbow angle. The mid upper arm circumference was measured at the midpoint of the upper arm length with the arm fully extended on the side, while maintaining the standing position. The circumference was then measured by wrapping the measuring tape around the arm (without compressing the skin and the underlying subcutaneous tissue) and keeping it perpendicular to the long axis of the upper arm.

During the leg length measurement, participants were first asked to sit on a chair while maintaining a 90° knee angle and a straight back. Leg length was measured as the length from the inguinal crease to the proximal border of the patella. During the thigh circumference measurement, the participants were asked to stand up and with their weight shifted on the left leg, lift their right leg off the ground with the knee slightly flexed. The circumference was then measured by wrapping the measuring tape around the midpoint of the upper leg, perpendicular to the long axis of the thigh (without compressing the skin).

During the calf circumference measurement, the participants were asked to maintain the same sitting position as in the leg length measurement. Calf circumference was measured by wrapping a measuring tape around the calf perpendicular to the long axis of the shank, at the site of maximal calf circumference. The site was found by moving the measuring tape up and down until the maximum calf circumference was detected. During the waist circumference measurement, the participants were asked to stand up, maintain a straight back, and extend their arms away from their trunk. The circumference was then measured by wrapping the measuring tape above the uppermost lateral border of the right ilium and around the mid-axillary line of the body.

Holtain skinfold calipers were used to measure the triceps and subscapular skinfold thickness to the nearest 0.1 mm. During the measurement, the participants maintained a standing position (with a straight back) with their arms extended away from the trunk. The measurement site for triceps thickness was the same as that for arm circumference measurement; that for subscapular thickness was at the inferior angle of the scapula. During the skinfold thickness measurement, the health technician first gently grasped the fold of the skin and underlying subcutaneous adipose tissue between the left thumb and index finger, and then grabbed the skin 2.0 cm above the measurement site to form a distinct fold that separates the skin from the muscle underneath. The skinfold for the triceps measurement was taken parallel to the long axis of the arm. The skinfold for the scapular measurement was taken so that the skinfold formed a line approximately 45° toward the right elbow. The thickness was then measured by placing the jaws of the calipers perpendicular to the length of the fold. A more detailed explanation of the protocols and sample videos of anthropometric measurements can be found on the NHANES website [18].

2.2.3. Body composition measures

A Hologic QDR 4500A fan beam X-ray bone densitometer (Hologic, Inc., Bedford, Massachusetts, USA) was used to perform whole-body DXA inside a mobile examination center by trained technologists. Before the scan, participants were asked to remove any jewelry they had worn and to lay on their backs on the DXA scan table. During the scan, the participants were asked to position themselves at the center of the table with their legs pointing inward and their toes touching. The technologist then tied a Velcro strap around the ankle to reduce movement. They were then asked to place their arms straight on their sides with palms facing downward, without touching the thighs. A single scan took approximately 3 min to complete. Five imputations were performed for each participant to acquire a more accurate variance estimate of body composition. The Hologic Discovery software (version 12.1) was used to calculate the regional and whole-body compositions, lean muscle mass, fat mass, and bone mineral density. We used whole-body lean muscle and fat mass for this study. The NCHS field staff monitored the scans performed by the technologists in the field, and the NHANES quality control center at the University of California checked the scan quality. A more detailed explanation of the DXA device, scan protocols, and steps taken to ensure quality control can be found on the NHANES website [18].

2.3. Development of lean muscle and fat mass estimation models

We built lean muscle and fat mass estimation models for datasets of ten categories of participants: all participants, separately for male and female participants, elderly, patients, and separately for five ethnic groups (Mexican American, Hispanic, White, Black, and Other). Elderly are selected as participants older than 65 years. Patients are selected as participants with any of the nine underlying diseases that impair the cardiovascular and metabolic functions of the participants (Anemia, Arthritis, Cancer, Congestive heart failure, Coronary heart disease, Thyroid, Liver diseases, and Chronic bronchitis). The presence of any of these diseases was determined using a survey assessing the medical conditions of the participants. For the dataset of all participants, elderly, and patients, 13 demographic and anthropometric measurement datasets were used as the input variables to estimate the two output variables separately: lean muscle and fat mass. Two demographic variables, sex and race, were treated as ordinal variables, whereas age, all anthropometric measures, and body composition measures were treated as continuous variables. For the models built separately according to sex and ethnicity, the number of input variables was reduced to 12 by removing either sex or ethnicity. Using the min–max normalization method, we rescaled all data to be between 0 and 1 before building the model.

First, we divided the dataset into training and validation sets. The training set was used to build a model that fits the measures to estimate lean muscle and fat mass, and the validation set was used as an independent set to evaluate the goodness of fit of the model. Consulting a previous study, we determined the ratio of training to validation set as 70:30 [17]. The training and validation sets were divided randomly 100 times to evaluate any effect of data division on the goodness of fit of each model. We evaluated the goodness of fit of the model using the coefficient of determination (R^2) and standard error of estimate (*SEE*) for both the training and validation sets. Following the NHANES guidelines on the analysis of multiple imputations, the goodness of fit of each model was averaged for the five imputations.

2.3.1. ANN model

We built a feed-forward, back-propagation ANN model comprising three layers (input, hidden, and output) for each of the five imputations. There were 13 nodes in the input layer for the dataset of all participants, elderly, and patients, one node for each demographic and anthropometric measure, and one node in the output layer for either lean muscle or fat mass. For the sex- or ethnicity-separated dataset, 12 nodes were used in the input layer. The transfer function used for the hidden layer was based on the Levenberg–Marquardt algorithm with Bayesian regularization [28,29] and that used for the output layer was a linear function. The algorithm used in the hidden layer served to determine the minimum mean square error between the estimated and actual lean muscle and fat mass, and to update the weight and bias values to improve the model's generalizability. The maximum number of iterations for the algorithm was set to 1000.

To determine the proper number of nodes in the hidden layer, we evaluated the *SEE* of the validation dataset by increasing the number of nodes from 1 to 25. For each set of nodes, we repeated the analysis 100 times by randomly dividing the training and validation datasets. We then averaged the *SEE* over the five imputations. The averaged *SEE* for a specific number of nodes was statistically compared with that for the next number of nodes. We found that the number of nodes for the hidden layer does not induce statistically significant differences between the neighboring pairs of node numbers when the node number increases beyond a specific value. We selected this specific value as the proper number of hidden layer nodes. For the dataset of all participants, the node numbers were selected as 6 and 7 for the models to estimate lean muscle and fat mass, respectively. For the dataset of male participants, we selected the number of hidden layer nodes as 5 for both lean muscle and fat mass estimation models. For the dataset of female participants, the node numbers were 4 and 3 for the models to estimate lean muscle and fat mass, respectively. For the elderly, the node numbers were 2 for both lean muscle and fat mass estimation models, whereas that for patients was 3 for both lean muscle and fat estimation models. For participants separated according to ethnicity, the number of nodes was selected as 3, 2, 5, 2, and 2 for lean mass estimation models for Mexican American, Hispanic, White, Black, and Other ethnic groups participants, respectively. For fat mass estimation models, the number of nodes was selected as 4, 3, 3, 4, and 3 for Mexican American, Hispanic, White, Black, and Other ethnic groups, respectively. A more detailed explanation of the procedure for selecting the number of nodes for the hidden layer is provided in the supplementary material.

2.3.2. MLR model

We built MLR models that estimated the dependent variables of lean muscle and fat mass using the independent variables of 13 demographic and anthropometric measures. For each of the sex- or ethnicity-specific models, the number of independent variables was set as 12 by removing either sex or ethnicity variable. The MLR models were built for the same 100 randomly divided training and validation datasets used for the ANN model, with the selected hidden layer nodes for each dataset category.

2.4. Statistical analysis

We performed paired t-tests to compare the *SEE* and R^2 of the ANN and MLR models. The 100 randomly divided sets made the sample size of each group 100. The paired t-test was conducted separately for the training and validation sets. The level of statistical significance (p) was set at <0.05. We also calculated the percentage difference (Δ) in the goodness of fit (*SEE* and R^2) values between the ANN and MLR models with respect to the baseline goodness of fit values of the MLR models.

3. Results

The sample size, demographics, anthropometry, and body composition of all ten categories of participants are summarized in Table 1. Figs. 1–6 show the mean and standard deviation values of the *SEE* and R^2 of 100 repetitions of randomly divided training and validation sets. The R^2 values of 100 repetitions were all above 0.92 for both training and validation sets of the ten categories of participants when ANN models were used, whereas the R^2 values were above 0.92 for all cases except the validation sets of elderly and patients for fat mass estimation when MLR models were used. The high R^2 values indicate that both ANN and MLR models explain the variability of a large portion of the dataset.

However, a notable difference exists between the two models. In every case, the *SEE* calculated using ANN models was lower than 2 kg for both the training and validation sets for all cases except for both sets of lean muscle and fat estimation models of elderly and patients, and both sets of fat and muscle estimation models for White and Black participants, respectively. On the other hand, the *SEE* calculated using MLR models was above 2 kg for all cases except in the case of validation sets for female participants. The paired t-tests revealed that ANN models always yielded lower *SEE* ($p < 0.001$) and higher R^2 ($p < 0.001$) values than the MLR models for both the training and validation sets, regardless of whether the datasets contained all participants, male, female, elderly, patients, or only a specific ethnic groups. The better performance of the ANN models was additionally quantified by the percentage difference (Δ) in the goodness of fit (*SEE* and R^2), which are shown in Figs. 1–6.

4. Discussion

Accurate measurement of lean muscle and fat mass in the human body is essential for the assessment of physical function and risk of cardiovascular and metabolic diseases, particularly in elderly and obese individuals. We developed and validated the demographic and anthropometric measure-driven ANN models to estimate lean muscle and fat mass using the NHANES dataset. Our models estimated these compartments of body compositions with a high goodness of fit, even when we separated the dataset according to sex or ethnicity. In addition, considering that body composition analysis is particularly important in patients and older adults, we further assessed the performance of the ANN models for those groups and confirmed that the ANN models still perform better than MLR models for such populations. These results confirm the efficacy and generalizability of the ANN models that estimate lean muscle and fat mass with simple inputs of easily obtainable demographics and anthropometric data.

The improved goodness of fit of ANN models over regression models has been observed in multiple previous studies. ANN models that estimate clinical outcomes such as mortality after traumatic brain injury [24], lung injury [30], and coronary artery diseases [25] showed higher accuracy than regression models. However, for the purpose of estimating lean muscle and fat mass and relating each mass to the risk of diseases and clinical outcomes caused by sarcopenia and obesity, previous studies depended on the MLR models developed by Lee et al. [17]. The MLR models were used to predict lean muscle and fat mass for independent populations and relate the mass to the risk of cardiovascular events [31], type-2 diabetes [32], and obesity-induced mortality hazard [33]. These studies showed that the estimated lean muscle and fat mass are good biomarkers for evaluating the risk of related diseases in a specific population. The demonstrated high performance of the ANN models developed in this study suggests that the models can assess the risk of body composition-related physiological deficiencies and the onset of metabolic diseases more reliably than previous MLR models.

A few studies have developed MLR models that estimate muscle and fat mass with similar goodness of fit to that of our models. Kulkarni et al. developed an anthropometry measures-driven MLR model using the data for Indian adults to estimate lean muscle and fat mass with R^2 values between 0.90 and 0.94 and *SEE* values between 1.47 and 1.92 kg [13]. Similar results were observed for models developed using the data of Chinese [14] and Japanese [15] adults with R^2 values between 0.81 and 0.93 and *SEE* values between 1.02 and 1.75 kg. The *SEE* values reported in these studies are lower than those reported in our models. However, the population groups used in these previous studies were mono-ethnic and had a small sample size; the lower *SEE* values can be partly due to the smaller range of lean muscle and fat mass values. The dataset used in this study contains data from a multi-ethnic population with a wide range of ages, anthropometric data, and lean muscle and fat mass values (Table 1). This diverse and large sample enhances both the generalizability and relative accuracy of the model; a larger sample size for training an ANN model can significantly improve the accuracy of output variable estimation [34]. Despite the wide range of output variables, the absolute *SEE* values of our models are comparable to those of previous models that focused on small samples of specific ethnic groups.

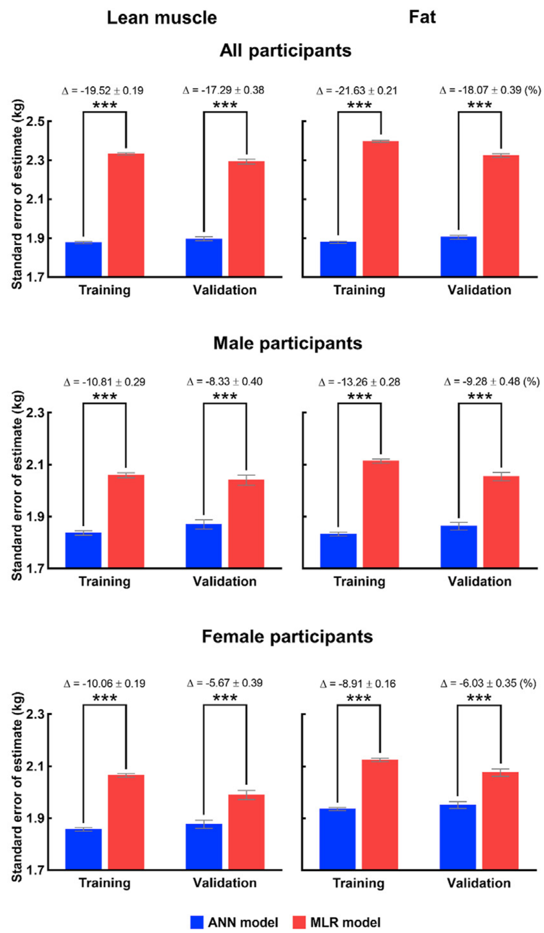
We observed that the improvement in the R^2 values is always smaller than that in the *SEE* values. Considering that the upper limit of R^2 values is 1, and the R^2 values are already high (above 0.92 for all cases) even for the conventional MLR models, any considerable improvements in R^2 values would be challenging. On the other hand, the *SEE* values have a relatively larger room for improvement, making it a more reasonable indicator to evaluate any improvement in model performance. Several studies have also suggested that the coefficient of determination may be an inadequate indicator of the performance of nonlinear and machine learning models [35–37].

We used 12 or 13 demographic and anthropometric measures to estimate lean muscle and fat mass, respectively. Previous studies have proposed models that predict the same variables with fewer anthropometric measures [8,12,17]. They simplified the model by reducing the independent variables to age, race, height, and weight, with the cost of increased *SEE* and decreased R^2 values. Although reducing the number of input parameters would reduce the time required for anthropometric measurement and model computation, the cost of the simplification, which is the deterioration in prediction performance, is critical. We additionally confirmed this

Table 1
Mean and standard deviations of the values of the demographic, anthropometric, and body composition measures.

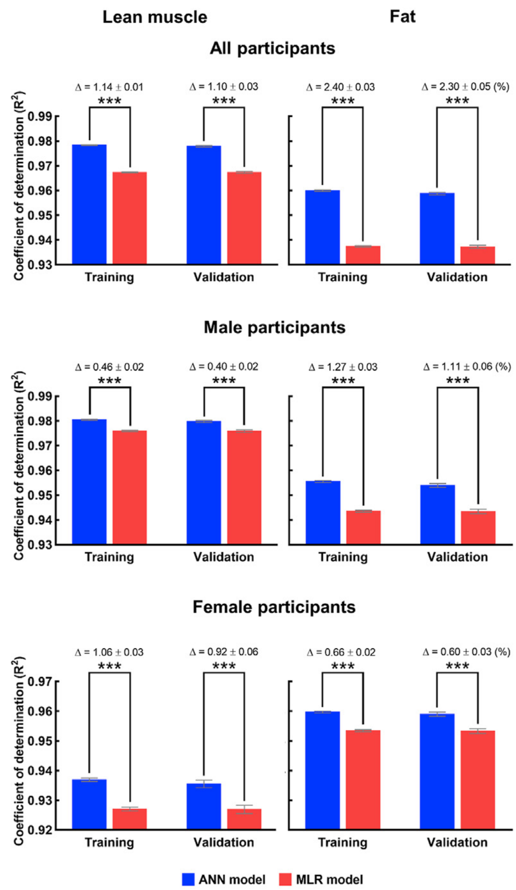
Measures		All	Male	Female	Elderly	Patients	Ethnicity				
							Mexican American	Hispanic	White	Black	Other
		n = 20,137	n = 11,319	n = 8818	n = 2358	n = 3414	n = 5462	n = 804	n = 8271	n = 4791	n = 809
Age (years)		33.033 ± 21.807	32.053 ± 21.392	34.290 ± 22.265	73.755 ± 5.505	60.237 ± 15.715	28.943 ± 20.277	31.909 ± 19.972	39.687 ± 22.719	26.903 ± 19.256	30.023 ± 19.868
Weight (kg)		67.264 ± 19.395	71.137 ± 20.558	62.292 ± 16.511	73.938 ± 14.875	76.542 ± 16.546	63.947 ± 18.413	65.370 ± 17.024	70.892 ± 19.125	65.947 ± 20.202	62.241 ± 19.586
Height (cm)		164.435 ± 13.259	168.946 ± 13.831	158.644 ± 9.814	165.244 ± 9.864	167.075 ± 10.024	160.376 ± 12.611	161.832 ± 12.122	167.450 ± 12.480	164.762 ± 14.000	161.666 ± 13.594
Upper leg length (cm)		39.908 ± 4.264	41.216 ± 4.279	38.170 ± 3.557	38.981 ± 3.800	39.397 ± 3.784	38.581 ± 4.158	39.169 ± 3.991	40.246 ± 3.957	41.118 ± 4.476	38.975 ± 4.273
Upper arm length (cm)		36.053 ± 3.567	37.083 ± 3.742	34.730 ± 2.822	37.272 ± 2.684	37.412 ± 2.792	35.100 ± 3.378	35.388 ± 3.218	36.730 ± 3.389	36.260 ± 3.816	34.989 ± 3.623
Arm circumference (cm)		29.532 ± 5.046	30.271 ± 5.136	28.584 ± 4.762	31.097 ± 3.712	31.745 ± 4.201	28.973 ± 4.932	29.551 ± 4.696	30.237 ± 4.841	29.131 ± 5.414	28.462 ± 5.018
Waist circumference (cm)		85.612 ± 15.496	87.067 ± 16.201	83.774 ± 14.326	97.888 ± 12.106	97.136 ± 13.283	85.434 ± 14.849	84.962 ± 13.559	88.858 ± 15.358	80.923 ± 15.501	82.034 ± 14.844
Thigh circumference (cm)		49.907 ± 6.901	50.300 ± 6.848	49.404 ± 6.937	49.061 ± 5.475	50.838 ± 6.036	48.596 ± 6.668	49.884 ± 6.676	50.313 ± 6.441	50.959 ± 7.632	48.404 ± 7.067
Calf circumference (cm)		36.093 ± 4.464	36.450 ± 4.523	35.634 ± 4.346	36.497 ± 3.589	37.446 ± 3.839	35.073 ± 4.329	35.926 ± 4.245	36.933 ± 4.292	35.926 ± 4.245	35.439 ± 4.627
Triceps skinfold (mm)		16.411 ± 7.856	13.056 ± 6.262	20.717 ± 7.591	17.847 ± 7.416	19.477 ± 8.207	16.252 ± 7.400	16.909 ± 7.990	17.214 ± 7.816	15.207 ± 8.308	15.900 ± 7.361
Subscapular skinfold (mm)		16.706 ± 8.052	15.593 ± 7.745	18.134 ± 8.213	19.301 ± 6.973	20.483 ± 7.615	16.860 ± 7.767	17.494 ± 7.898	17.014 ± 7.856	15.877 ± 8.544	16.633 ± 8.105
Lean muscle mass (kg)	**1**	44.762 ± 12.899	50.448 ± 13.287	37.464 ± 7.657	45.857 ± 10.216	47.588 ± 11.412	42.017 ± 12.107	43.231 ± 11.694	46.542 ± 12.805	45.580 ± 13.446	41.781 ± 13.108
	2	44.763 ± 12.903	50.453 ± 13.291	37.460 ± 7.651	45.863 ± 10.218	47.605 ± 11.442	42.021 ± 12.113	43.226 ± 11.683	46.543 ± 12.811	45.578 ± 13.446	41.779 ± 13.097
	3	44.764 ± 12.901	50.450 ± 13.293	37.466 ± 7.651	45.882 ± 10.219	47.625 ± 11.449	42.023 ± 12.110	43.230 ± 11.678	46.543 ± 12.805	45.578 ± 13.452	41.793 ± 13.112
	4	44.760 ± 12.900	50.447 ± 13.289	37.459 ± 7.650	45.850 ± 10.207	47.612 ± 11.446	42.019 ± 12.109	43.222 ± 11.681	46.541 ± 12.807	45.570 ± 13.444	41.781 ± 13.109
	5	44.762 ± 12.900	50.452 ± 13.288	37.459 ± 7.646	45.870 ± 10.222	47.615 ± 11.426	42.024 ± 12.109	43.216 ± 11.681	46.543 ± 12.807	45.575 ± 13.440	41.769 ± 13.109
Fat mass (kg)	**1**	20.972 ± 9.586	19.100 ± 8.898	23.493 ± 9.845	26.426 ± 8.262	27.217 ± 9.092	20.561 ± 8.968	20.718 ± 8.824	22.740 ± 9.522	18.760 ± 10.034	19.046 ± 8.865
	2	20.970 ± 9.583	19.005 ± 8.892	23.493 ± 9.844	26.416 ± 8.257	27.212 ± 9.070	20.554 ± 8.957	20.718 ± 8.819	22.737 ± 9.518	18.760 ± 10.039	19.055 ± 8.870
	3	20.972 ± 9.584	19.012 ± 8.896	23.487 ± 9.845	26.404 ± 8.258	27.193 ± 9.043	20.557 ± 8.960	20.720 ± 8.825	22.741 ± 9.521	18.761 ± 10.034	19.032 ± 8.855
	4	20.975 ± 9.585	19.012 ± 9.898	23.494 ± 9.843	26.445 ± 8.266	27.202 ± 9.043	20.558 ± 8.962	20.724 ± 8.816	22.742 ± 9.518	18.768 ± 10.046	19.039 ± 8.847
	5	20.975 ± 9.590	19.009 ± 8.895	23.499 ± 9.856	26.423 ± 8.276	27.206 ± 9.063	20.554 ± 8.961	20.728 ± 8.827	22.742 ± 9.524	18.769 ± 10.051	19.060 ± 8.881
Ethnicity	**Mexican American (%)**	27.124	27.856	26.185	18.999	13.533					
	Hispanic (%)	3.993	3.808	4.230	3.053	1.904					
	White (%)	41.074	39.544	43.037	62.595	65.788					
	Black (%)	23.792	24.949	22.307	13.062	14.968					
	Other (%)	4.017	3.843	4.241	2.290	3.808					

Note: The number assigned from 1~5 for the lean muscle and fat mass refers to the imputation number. For ethnicity, data are presented in terms of the percentage of total sample size.



*** (p<0.001) represents significant differences in standard error of estimate between the ANN and MLR models

Fig. 1. Mean and standard deviation of the standard error of estimate (*SEE*) values of 100 randomly divided repetitions for training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) for all, male, and female participants. Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the *SEE* values between the ANN and MLR models with respect to the baseline *SEE* values of MLR model.



*** (p<0.001) represents significant differences in coefficient of determination (R^2) between the ANN and MLR models

Fig. 2. Mean and standard deviation of the coefficient of determination (R^2) values of 100 randomly divided repetitions of training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) for all, male, and female participants. Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the R^2 values between the ANN and MLR models with respect to the baseline R^2 values of MLR model.

by building simplified versions of the ANN and MLR models using only height and weight as anthropometric measures for all participants' datasets (supplementary material). As with the results of previous studies, reducing the anthropometric measures resulted in a significant reduction in the goodness of fit. On the other hand, instruments such as stadiometers, digital weight scales, measurement tapes, and skin calipers, used for anthropometric measurements are inexpensive and readily available in various areas lacking expensive infrastructure and equipment for body composition measurement. The process of anthropometric measurement is also straightforward. Hence, if we prioritize finding an inexpensive but reliable alternative to DXA or MRI in estimating lean muscle and fat mass, it would be beneficial to spend a short additional time measuring limb circumferences and lengths to secure the reliability of the estimation model.

Previous studies have reported that modulation of body composition significantly impacts the survival probability and mortality [38–40]. Demographic and anthropometric measures-driven MLR models developed by Lee et al. [17] have been extensively used in multiple epidemiological studies to estimate body composition and establish its relationship with mortality of large populations [31,33,41]. The same group of researchers have also highlighted that the measurement error is a critical factor in establishing an accurate relationship between body composition and mortality [42]. Therefore, considering the improved accuracy of the ANN model we developed, we postulate that using body composition estimated by the ANN models would improve the accuracy when establishing an association between body composition and mortality in epidemiological studies. However, the current dataset does not allow directly comparison between ANN and MLR models regarding the accuracy in estimating mortality because NHANES is not a cohort study that regularly measures demographic, anthropometric, and body composition measures and mortality. The records of long-term changes in body composition and mortality are necessary to properly evaluate the association between changes in body composition and mortality. A future work of extensive cohort study with regular measurements of demographic, anthropometric, body composition measures, and mortality will enable researchers to assess the performance of ANN models in predicting body composition-related modulation of mortality.

Although the ANN models showed better performance than MLR models, ANN models are inherently complex, frequently prone to overfitting, and require larger computational power and

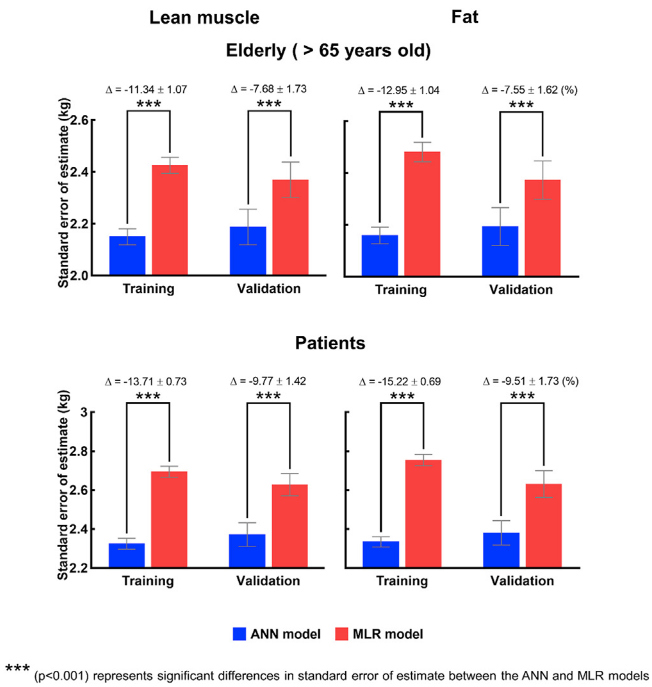


Fig. 3. Mean and standard deviation of the standard error of estimate (*SEE*) values of 100 randomly divided repetitions of training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) for elderly and patients. Elderly are categorized as participants older than 65 years and patients are categorized as participants with any underlying illness. Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the *SEE* values between the ANN and MLR models with respect to the baseline *SEE* values of MLR model.

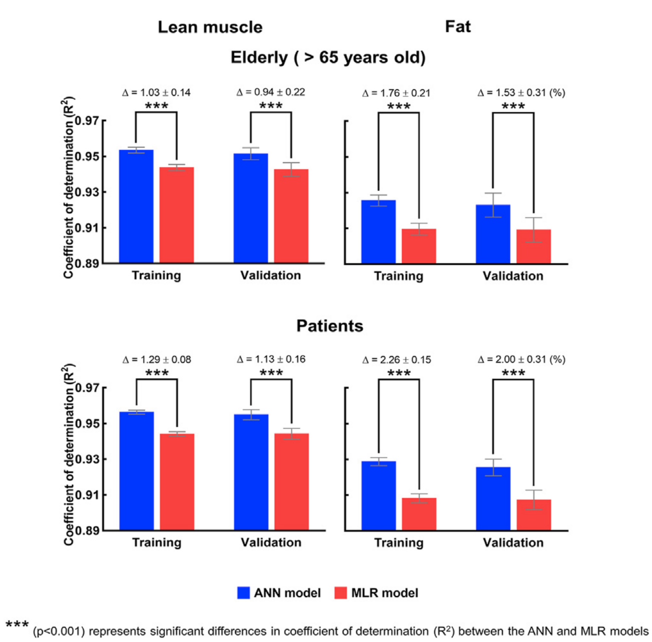


Fig. 4. Mean and standard deviation of the coefficient of determination (R^2) values of 100 randomly divided repetitions of training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) for elderly and patients. Elderly are categorized as participants older than 65 years and patients are categorized as participants with any underlying illness. Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the R^2 values between the ANN and MLR models with respect to the baseline R^2 values of MLR model.

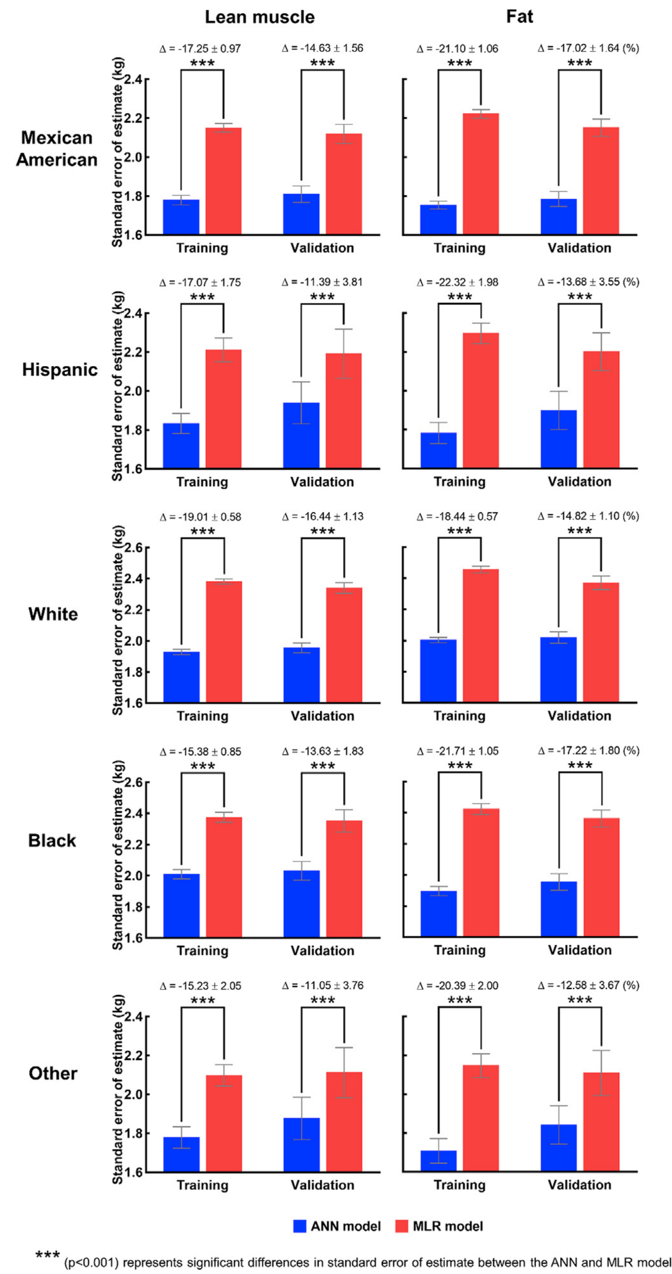


Fig. 5. Mean and standard deviation of the standard error of estimate (*SEE*) values of 100 randomly divided repetitions of training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) separately for five ethnicities (Mexican American, Hispanic, White, Black, and Other). Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the *SEE* values between the ANN and MLR models with respect to the baseline *SEE* values of MLR model.

processing time. Fortunately, current advancements in computing technology are sufficient for simple single-layered ANN models (such as those used in our study) to estimate output variables swiftly; the computation time required to build an ANN model to estimate muscle and fat mass estimation using the largest dataset (all participants) was between 7 and 8 s. We also pruned the hidden layer parameters to determine the minimal number of nodes without significantly compromising the estimation errors of the model (supplementary material). This pruning process contributed

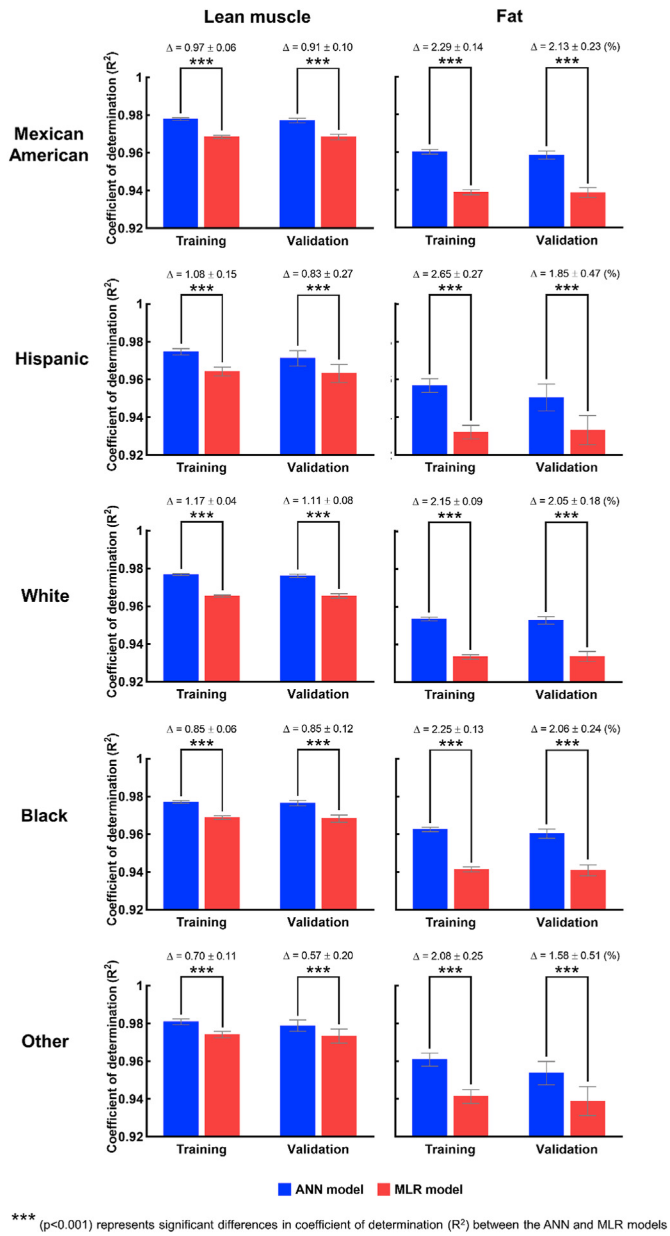


Fig. 6. Mean and standard deviation of the coefficient of determination (R^2) values of 100 randomly divided repetitions of training and validation sets for estimation models developed using artificial neural network (ANN) and multiple linear regression (MLR) separately for five ethnicities (Mexican American, Hispanic, White, Black, and Other). Among the total dataset, the training and validation sets were divided into a 70:30 ratio. The bars denote standard deviations. Δ denotes percentage differences in the R^2 values between the ANN and MLR models with respect to the baseline R^2 values of MLR model.

to reducing the complexity of the model and minimizing overfitting. This process and its result can serve as a reference when other researchers set out to develop other ANN models using NHANES data and need to decide the proper number of nodes.

5. Conclusion

We offer a cost-effective and straightforward demographic and anthropometric-measure-driven ANN model that can estimate lean muscle and fat mass with high precision. Although our method is more complex than that adopted in previous studies, the resulting estimation model significantly reduces prediction errors. We also outline in detail the process of developing the ANN models. This process can be implemented by researchers and medical professionals for their datasets to build competent models that estimate body composition using easily measurable parameters. We expect that our method and results can be used in large population epidemiological studies and contribute to addressing body composition-related issues using minimal resources.

Funding

This work was supported in part by National Research Foundation of Korea (NRF) Grants funded by the Korean Government (MSIT) (No. 2016R1A5A1938472), the Technology Innovation Program (No. 20008912) and Industrial Technology Innovation Program (No. 20007058, Development of safe and comfortable human augmentation hybrid robot suit) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea), and Ministry of Education of the Republic of Korea.

Authors' contributions

Conceptualization: PP; Data curation: PP and SBP; Formal analysis: PP; Funding acquisition: JA; Methodology: PP and SBP; Investigation: PP and SBP; Project Administration: JA; Resources: PP and SBP; Software: PP; Supervision: JA; Validation: PP, SBP, and JA; Visualization: PP; Roles/Writing – original draft: PP; Writing – review & editing: JA.

Conflict of interest

The authors declare no competing interests.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clnu.2021.11.027.

References

[1] Pietrobelli A, Heymsfield S, Wang Z, Gallagher D. Multi-component body composition models: recent advances and future directions. Eur J Clin Nutr 2001;55:69–75.
[2] Caballero B, Trugo LC, Finglas PM. Encyclopedia of food sciences and nutrition. Academic; 2003.
[3] Jiménez EG. Body composition: assessment and clinical value. Endocrinol Diabetes Nutr 2013;60:69–75.
[4] Adams KF, Schatzkin A, Harris TB, Kipnis V, Mouw T, Ballard-Barbash R, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. N Engl J Med 2006;355:763–78.
[5] Visser M, Schaap LA. Consequences of sarcopenia. Clin Geriatr Med 2011;27:387–99.
[6] Oreopoulos A, Kalantar-Zadeh K, Sharma AM, Fonarow GC. The obesity paradox in the elderly: potential mechanisms and clinical implications. Clin Geriatr Med 2009;25:643–59.
[7] Duren DL, Sherwood RJ, Czerwinski SA, Lee M, Choh AC, Siervogel RM, et al. Body composition methods: comparisons and interpretation. J Diabetes Sci Technol 2008;2:1139–46.
[8] Ross R, Shaw KD, Rissanen J, Martel Y, de Guise J, Avruch L. Sex differences in lean and adipose tissue distribution by magnetic resonance imaging: anthropometric relationships. Am J Clin Nutr 1994;59:1277–85.
[9] Pereira PMG, da Silva GA, Santos GM, Petroski EL, Geraldes AAR. Development and validation of anthropometric equations to estimate appendicular muscle mass in elderly women. Nutr J 2013;12:1–11.
[10] Poortmans JR, Boisseau N, Moraine J-J, Moreno-Reyes R, Goldman S. Estimation of total-body skeletal muscle mass in children and adolescents. Med Sci Sports Exerc 2005;37:316–22.

[11] Rech CR, Dellagrana RA, Marucci MdFN, Petroski EL. Validity of anthropometric equations for the estimation of muscle mass in the elderly. Rev Bras de Cineantropometria e Desempenho Hum 2012;14:23–31.
[12] Lee RC, Wang Z, Heo M, Ross R, Janssen I, Heymsfield SB. Total-body skeletal muscle mass: development and cross-validation of anthropometric prediction models. Am J Clin Nutr 2000;72:796–803.
[13] Kulkarni B, Kuper H, Taylor A, Wells JC, Radhakrishna K, Kinra S, et al. Development and validation of anthropometric prediction equations for estimation of lean body mass and appendicular lean soft tissue in Indian men and women. J Appl Physiol 2013;115:1156–62.
[14] Xu W, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20:551–6.
[15] Furushima T, Miyachi M, Iemitsu M, Murakami H, Kawano H, Gando Y, et al. Development of prediction equations for estimating appendicular skeletal muscle mass in Japanese men and women. J Physiol Anthropol 2017;36:1–9.
[16] Sesbreno E, Slater G, Mountjoy M, Galloway SD. Development of an anthropometric prediction model for fat-free mass and muscle mass in elite athletes. Int J Sport Nutr Exerc Metabol 2020;30:174–81.
[17] Lee DH, Keum N, Hu FB, Orav EJ, Rimm EB, Sun Q, et al. Development and validation of anthropometric prediction equations for lean body mass, fat mass and percent fat in adults using the National Health and Nutrition Examination Survey (NHANES) 1999–2006. Br J Nutr 2017;118:858–66.
[18] CDC. National health and nutrition examination survey.
[19] Dayhoff JE, DeLeo JM. Artificial neural networks: opening the black box. Cancer 2001;91:1615–35.
[20] Schmidhuber J. Deep learning in neural networks: an overview. Neural Network 2015;61:85–117.
[21] Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49:1225–31.
[22] Ruiz JR, Ramirez-Lechuga J, Ortega FB, Castro-Pinero J, Benitez JM, Arauzo-Azofra A, et al. Artificial neural network-based equation for estimating VO2max from the 20 m shuttle run test in adolescents. Artif Intell Med 2008;44:233–45.
[23] Disse E, Ledoux S, Bétry C, Caussy C, Maitrepierre C, Coupaye M, et al. An artificial neural network to predict resting energy expenditure in obesity. Clin Nutr 2018;37:1661–9.
[24] Shi HY, Hwang SL, Lee KT, Lin CL. In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. J Neurosurg 2013;118:746–52.
[25] Akella A, Akella S. Machine learning algorithms for predicting coronary artery disease: efforts toward an open source solution. Future Sci OA 2021:FSO698.
[26] CDC. National health and nutrition examination survey: technical documentation for the 1999-2004 dual energy X-ray absorptiometry (DXA) multiple imputation data files. 2008.
[27] Lohman TG, Roche AF, Martorell R. Anthropometric standardization reference manual. Human kinetics books Champaign; 1988.
[28] MacKay DJ. Bayesian interpolation. Neural Comput 1992;4:415–47.
[29] Foresee FD, Hagan MT. Gauss-Newton approximation to Bayesian learning. Proceedings of international conference on neural networks (ICNN'97). IEEE; 1997. p. 1930–5.
[30] Fei Y, Gao K, Li WQ. Artificial neural network algorithm model as powerful tool to predict acute lung injury following to severe acute pancreatitis. Pancreatology 2018;18:892–9.
[31] Xing Z, Tang L, Chen J, Pei J, Chen P, Fang Z, et al. Association of predicted lean body mass and fat mass with cardiovascular events in patients with type 2 diabetes mellitus. Can Med Assoc J 2019;191:E1042–8.
[32] Lee DH, Keum N, Hu FB, Orav EJ, Rimm EB, Willett WC, et al. Comparison of the association of predicted fat mass, body mass index, and other obesity indicators with type 2 diabetes risk: two large prospective studies in US men and women. Eur J Epidemiol 2018;33:1113–23.
[33] Lee DH, Keum N, Hu FB, Orav EJ, Rimm EB, Willett WC, et al. Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: prospective US cohort study. Br Med J 2018;362.
[34] Riley RD, Ensor J, Snell KI, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. Br Med J 2020:368.
[35] Kvålseth TO. Cautionary note about R2. Am Statistician 1985;39:279–85.
[36] Anderson-Sprecher R. Model comparisons and R 2. Am Statistician 1994;48:113–7.
[37] Spiess A-N, Neumeyer N. An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. BMC Pharmacol 2010;10:1–11.
[38] Heitmann B, Erikson H, Ellsinger B, Mikkelsen K, Larsson B. Mortality associated with body fat, fat-free mass and body mass index among 60-year-old Swedish men—a 22-year follow-up. The study of men born in 1913. Int J Obes 2000;24:33–7.
[39] Schols AM, Broekhuizen R, Weling-Scheepers CA, Wouters EF. Body composition and mortality in chronic obstructive pulmonary disease. Am J Clin Nutr 2005;82:53–9.
[40] Wannamethee SG, Shaper AG, Lennon L, Whincup PH. Decreased muscle mass and increased central adiposity are independently related to mortality in older men. Am J Clin Nutr 2007;86:1339–46.
[41] Jeong S-M, Lee DH, Giovannucci EL. Predicted lean body mass, fat mass and risk of lung cancer: prospective US cohort study. Eur J Epidemiol 2019;34:1151–60.
[42] Lee DH, Giovannucci EL. Body composition and mortality in the general population: a review of epidemiologic studies. Exp Biol Med 2018;243:1275–85.