Research Article

Development of Body Fat Estimation Equations Based on BMI, Age, Gender, and Ethnic Groups

Hsuan-Yu Chen[1], Chiachung Chen[2]

1. Africa Industrial Research Center, National Taipei University, Taiwan, Province of China; 2. National Taipei University, Taiwan, Province of China

Body fat (BF) percentage is a measurement of human health. The simple predictive equation between BF and the anthropometric measurement helps evaluate the BF value. The BMI value, weight divided by the height square, has been used as the significant factor in BF values. However, other factors involving age, gender, and ethnicity may also affect the BF values. An adequate model considering all influencing factors is critical in predicting the BF value. Many empirical equations have been proposed to evaluate these factors. This study uses previously collected data to establish the BF equation with modern regression analysis. Three forms of body mass index (BMI), BMI, logarithmic BMI, and inverse BMI, are selected as independent valuables. The other variables include age, gender, and ethnicity. The t-value was used to test the significant influence of each variable in the regression equations. The Prediction Sum of Squares (PRESS) statistics were used to evaluate the models' predictive ability. Categorical testing was adopted to evaluate the significant influence of these variables of age, gender, and ethnicity. The results of this study indicated that the best BF model involved BMI, BMI2, age, and age2 variables. The age, gender and ethnicity tested by categorical test significantly affected BF values. No single form of the BF equation can be proposed to represent all ethnicities or both genders. Modern regression analysis can provide more scientifically based model-building techniques than machine learning. By calculating the BMI value, the cutoff point of BF values needs to consider the difference between gender and ethnicity. The regression technique in this study provides a reasonable method to establish the BF equation for different ethnicities and other factors.

Corresponding author: Chiachung Chen, ccchen@dragon.nchu.edu.tw

 doi.org/10.32388/T73GYR

1. Introduction

Body fat percentage (BF) is an essential measurement of human health. It also affects the pharmacokinetics of most drugs[1]. The fat fraction significantly affects the human propofol kinetics[2]. The BF values could be mesured directly by using bioimpedance spectroscoply (BIS), bioelectrical impedance analysis (BIA), air displacement plethysmography, and electrical impedance myography (BIM). However, these measurement methods are expensive and time-consuming. Body mass index (BMI) is calculated by a subject's weight and height and is then used to predict the BF value. It provides a simple way to find the BF values. Overweight or obesity is a risk factor for chronic disease. The cut-off points for overweight or obesity are generally defined as a certain BMI threshold value[3][4].

Some empirical models have been proposed to describe the relationship between BF and other influencing factors. Due to its simplicity, BMI serves as the standard variable to predicate BF and is then used to classify the degree of obesity. However, some evidence suggests the BMI cut-points may not be the same for all ethnic groups[5]. The relationship between fat and BMI in gender and different ethnic groups may provide an explanation.

Some researchers reported the change in fat at a given BMI is affected by age and gender[6][7]. The relationship between BF and BMI and other influencing factors needs to be further studied and validated.

Researchers have introduced different empirical equations. Deurenberg et al.[7] proposed two linear equations representing the effect factors on BF. The coefficient of determination, R2, and estimated value of standard errors of these models were used to evaluate the adequacy of the proposal model. The ethnicity of the subjects was not mentioned. Gallagher et al.[8] introduced the linear equation to express the relationship between BF and BMI. Age and gender did not significantly influence BF. However, Deurenberg et al.[9] proposed the predictive equation: BF = b0 +b1BMI+b2age+b3gender. The effect of ethnicity on the BF values was observed by the calculated values at fixed BMI values in subjects of different ethnicities, with the effects found to be inconsistent.

Gallagher et al.[10] introduced a complex multiple-linear equation to predict BF values. The independent variables included the inverse function of BMI (BMI-1), sex, age, ethnicity, and the interaction of variables. Jackson et al.[11] studied the relationship between BF and BMI for subjects

doi.org/10.32388/T73GYR

including race (white and black) and gender (men and women). The polynomial equation was established as BF=c0+c1BMI+c2BMI2. Age, ethnicity, and ethnicity and BMI interaction were found to have no significant effect on the BF value.

To establish the obesity cutoffs for Taiwanese subjects, Chang et al.[12] proposed a linear regression model as BF = d0 + d1BMI, and the new BMI cutoffs of overweight and obesity were suggested for male and female Taiwanese adults. Kagawa et al.[13] performed a study to evaluate the relationship between BF and BMI for young Japanese and Australian men. Both independent and dependent variables were in logarithmical form, LnBF= e0 + e1Ethnicity + e2LnBMI. Ethnicity significantly affected the relationship between LnBMI and LnBF. Larson et al.[14] measured BF and other anthropometric variables to assess the effect variables on the BF values. Two indexes, weight/height (W/H) and weight/(height)2, were compared for the predictive performance. Their result indicated that the W/H value may be a more valuable index than that of W/H2 (BMI) in expressing the effect on BT, especially at higher body weights. However, the criteria to evaluate the performance of these models were the correlation coefficient and absolute errors. Rush et al.[15] compared the difference between BMI and BF values in women of five ethnicities from two countries. The proposed equation was BF = f0 + f1LogBMI. The R2 and s were the criteria to evaluate these data. Their results indicated no single form of the BF equation and BMI cut-points could be established. Meeuwsen et al.[16] observed UK adults and proposed a curvilinear mode to express the influence factors on BF. Gomez-Ambrosi et al.[17] recommend a BF prediction equation with more variables and interaction of variables, such as age, gender, BMI, BMI2, BMI*gender, BMI*age, BMI2∗gender, BMI2∗age.

The model developed by Gomez-Ambrosi et al.[17] was adopted by Fuster-Parra et al.[18]. The fitting agreements of the BF model with BMI and the body adiposity index (BAI) were compared, and it was found that the BAI variable had better predictive ability than that of BMI. The correlation coefficient served as the criterion for comparing.

Cortes-Castell et al.[19] developed a new model to predict the BF values of children in Spain. Itani et al.[20] proposed the BF models for the mean and women of Labanese. The variables included BMI2, 1/BMI, and age. In research on body fat prediction models in American adults[21], more anthropometric measurements, such as waist circumference, hand thickness, vertical abdominal skinfold, and thigh skinfold, were incorporated into the BF model.

To consider obesity as a public health problem, the BF prediction model was developed in different countries, such as Mexico[22], Vietnam[23], Nigeria[24], Serbia[25], Saudi Arabia[26], Italy[27], Philippines[28], and Lebanon[20].

The published models using data involving several variables are listed in Table 1.

Reference	Regression equations	BMI range	
1. Womersley and Durnin[7]	1a. Men BF=-12.47+1.34BMI 1b. Women BF = -3.47+1.37BMI		
2.Deurenberg et al.[6]	2a. Children (ages $\leq$ 15) BF=1.4+1.51 BMI-0.70 age-3.6gender, R2=0.38, s=0.044 2b. Adults BF= -5.4+1.20BMI+0.23age-10.8gender R2=0.79, s=0.041	16-36.5	
3.Gallagber et al.[8]	3a. Black, Women BF=-6.254+1.419BMI+0.07age, R2=0.58 3b. White, Women BF=-11.666+0.1591BMI+0.096age, R2=0.56 3c. Black, Men BF=-18.624+1.367BMI+0.105age, R2-0.44 3d. White, Men BF=-22.519+1.402 BMI+0.177age, R2=0.52		
4.Deurenberg et al.[9]	Caucasian BF=-8.0+1.294BMI+0.20Age-11.4gender, R2=0.88	19-31	
5. Gallagher et al.[10]	White and African American BF=84.8/BMI+0.079age+16.4gender+0.05Sgender*age+39.0gender/BMI, R2=0.86	16-38	
6.Jackson et al.[11]	6a. Female BF=-82.83+34.43InBMI+0.14age-26.02gender-7.48gender*BMI, R2=0.82, 6b. Male BF=-149.24+51.31InBMI+1.47age-0.41age*InBMI, R2=0.82 6c. Heritage Woman BF=-46.24+4.35BMI-0.05BMI2, R2=0.78 6d. Heritage Man BF=-47.8+3.76BMI-0.04BMI2, R2=0.68	17-47	
7.Chang et al.[12]	Taiwan 7a. Women BF=-3.07+1.652BMI, R2=0.6	14-33~15-33	

doi.org/10.32388/T73GYR

Reference	Regression equations	BMI range	
	7b. Men BF=-9.093+1.38BMI, R2=0.475		
8. Kagawa et al. [13]	Japanese and Australian Caucasian InBF=-3.321-0.123Ethnicity+1.941In BMI, R2=0.548, s=0.21	15.6~36.6	
9. Larson et al. [14]	Swedish 9a. Women BF=-28.27+1.895BMI, R2=0.61 9b. Man BF=-49.47+3.865 BMI-0.0343 BMI2, R2=0.72	17-46	
10. Rush et al. [15]	Women, five ethnicities, two countries 10a. SA, European BF= -83.6+82.9 LogBMI, R2=0.73, s=4.9 10b. NZ, European BF=-111.2+105.2 LogBMI, R2=0.80, s=4.8	16.1~48.0	
11. Levitt et al. [29]	Caucasian, Hispanic, Black 11a. Women BF=64.2-9.23/BMI +0.132age 11b. Men BF=-20.6+1.27BMI+0.182age	17.07-65.75	
12.Meeuwsen et al.[16]	UK adults BF=-32.515+12.409gender+3.306BMI-0.030BMI2-0.006age +0.033age*gender-0.001age*BMI,	14.5-45.5	
13.Gomez-Ambrosi et al. [17]	Caucasian, Spain adult BF=-44.988+0.503age+10.689gender +3.172 BMI-0.026BMI2 +0.181•BMI∗gender -0.02BMI∗Age-0.005BMI2*gender +0.0021 BMI2*age, R2=0.774, s=4.7	13.8-70.2	
14. Fuster-Parra et al.[18]	Caucasian, Spain adults BF=-44.988+0.503age+10.689 gender+3.172BMI-0.026BMI2 +0.181BMI∗gender-0.02•BMI∗Age-0.005BMI2∗gender +0.00021•BMI2∗Age, R2=0.79	15.8~51.3	

doi.org/10.32388/T73GYR

Reference	Regression equations	BMI range	
15. Ho-Pham et al.[23]	Vietnamese BF=18.9-10.9gender+0.044age +3.472BMI-0.05BMI2 R2=0.71, s=3.90	15-36	
16. Akindele et al.[24]	Nigeria BF=0.583BMI+0.456gender+0.152age, R2=0.82	10-58	
17. Jelena et al.[25]	Serbian, adults BF=-5.4+0.23age-10.8gender+1.2BMI	16.8-33	
18. Eldeen et al.[26]	Saudi Arabian, adults BF=-5.4+0.23age-10.8gender+1.20BMI	15-30.5	
19. Cortes-Castell et al.[19]	Spain children 19a. Boy BF= 18.655+0.007BMI2-293.601/BMI+0.112age-0.018age2 19b. Girls BF=18.655+0.007BMI2-293.601/BMI	18.6-34.6	
20. Itani et al.[20]	Lebanese, adults 20a. Women BF= 21.835+0.622BMI, R2=0.718 s=3.81 20b. Men BF=-4.001+1.050 BMI, R2=0.709 s=3.88	24-55	
21. Molina-luque et al.[21]	Spain, adults BF=-97.102+0.123age-11.9gender+35.959BMI	21.5-30.8	
22. Chen et al.[30]	Singaporean, adults 19a. Men BF= 49.818+00.089age-610.808/BMI 19b. Women BF=58.159+0.051age-516.401/BMI	15.1-39.9	
23. Kalaragunta et al.[31]	India, adults 20a. Women BF= -5.4+0.23age-10.8gender+1.2BMI	18.5-29.9	

Table 1. Published models and Statistical tests of the body mass index (BMI) and body BF.

doi.org/10.32388/T73GYR

Note: BF: percentage body fat in %, BMI: body mass index in Kg/m2, age in years, gender: men and women, categorical variables.

Recently, researchers have examined the predictive performance of different BF models. Nickerson et al.[32] evaluated the relative accuracy of Womersley and Durnin[6], Jackson et al.[11], Deurenberrg et al.[7] and Gallagher et al.[10]. Inconsistent results were found in subjects of both men and women. Jaafar et al.[33] assessed the accuracy of the BF estimation in Arab people with the above four models and the Gomez-Ambrosi model[17]. The equation of Jackson et al.[11] was recommended as the best model for obesity in the Lebanese population. De Nortoli et al.[34] recommended that the Gomez-Ambrosi equation be better than other BMI equations for patients undergoing bariatric surgery.

In the above introduction, most studies use R2 as the sole criterion. Some studies select s and the p-value to evaluate the validations of a BF model. However, these criteria (R2, s, and p-value) are only used for classical regression[35][36].

Besides the regression methods, Faradisa et al.[37] used fuzzy logic to establish the relationship between BF and BMI. The fuzzy sets of body weight, height, and BMI were 3, 3, and 5, respectively. They found that the BF and BMI can produce the same categories. Xu et al.[38] adopted a supervised machine-learning approach to develop the prediction equation for BF and BMI. Gender, income, age, and education were the affecting factors. Their best model included 18 variables and 19 parameters.

This study adopts a modern regression technique to evaluate the adequateness of different BF models and determine the factors affecting the BF values. Modern regression techniques, such as residual plots, the normal test, the constant variance test, the test on a single regression coefficient, and categorical testing, are used in this study to develop the BF models. The statistics, Prediction Sum of Squares (PRESS) is used to assess the prediction ability of the models[28][32]. A modern regression technique was applied in liver volume prediction[33], plant tissue culture[34], water activity in floral honey[35], and dielectric properties of foods[36].

To the best of the authors' knowledge, the modern regression technique has not been adopted in the study of BF models. Here, the adequateness of BF models was evaluated, and the factors affecting the Bf values of different subjects were determined using data from previous studies.

2. Materials and Methods

2.1. Regression analysis

In this study, the dependent variable yi is the BF. The independent variables include BMI, age, gender, and ethnicity. The BMI and age are continuous variables, and gender and ethnicity are categorical variables.

The effect of BMI and age on BF is evaluated with a typical multiple regression model:

BF = b0 + b1BMI + b11BMI2 + b2age + b22age2 +b12BMI*age (1)

where b_0 , b_1, b_{11}, b_2, b_{22}, b_{12} are constants.

This equation is called the BMI equation.

The diagnostic techniques of modern regression analysis are described as follows:

2.1.1. Normality test

The Kolmogorov-Smirnov test verifies the normal data distribution. The p-value determines its probability.

2.1.2. Residual Plots

For the residual plot, the residuals are on the longitudinal axis, and the predicted value of the model is on the horizontal axis. If the uniform distribution of errors along the yi=0 line is presented, the regression model is adequate. If the fixed pattern of the data distribution is found, the model is not suitable. The funnel pattern indicates that the error variance is not constant; that is, it is a heterogeneous variance.

2.1.3. Transformation

Two forms of transformation are used in this study. Previous research has recommended this transformation[4][13][19]. Two equations are proposed for further study:

BF = c0 + c1LnBMI + c11(LnBMI)2 + c2age + c22age2 +c12(LnBMI)*age (2)

where c0, c1, c11, c2, c22 and c12 are constants.

The equation is called the logarithmic BMI equation.

doi.org/10.32388/T73GYR

BF = d0 + d1(1/BMI) + d11(1/BMI) 2 + d2age + d22age2 +d12(1/BMI) *age (3)

where d0, d1, d11, d2, d22 and d12 are constants.

The equation is called the inverse BMI equation or 1/BMI equation.

2.2. *Categorical testing*

Influencing factors, such as gender and ethnicity, are categorical variables. The categorical test technique is used to determine their significance.

2.2.1. *Two categories.*

The two categories of gender (men, women) and ethnicity (Black, Mexican) can be tested for their level of effect on the BF values. The BF equation linking the dependent variable and two variables is:

BF = b0 + b1BMI + b11BMI2 + b2age + b22age2 +b12BMI*age (4)

Let Z be defined as follows:

Z=0, if in the male category

Z=1, if in the female category

The equation to determine the significant effect of the two levels as a factor is:

$$BF = e_0 + e_3 Z + e_1 BMI + e_{13} BMI * Z + e_{11} BMI^2 + e_{113} BMI^2 * Z + e_2 age + e_{23}\ age * Z$$

$$+ e_{22} age^2 + e_{23} age^2 \bullet Z + e_{12} BMI * age + e_{123} BMI * age * Z + \varepsilon_i \ (5)$$

where all e_{ii} are constants, and Z is the categorical value. The significant effect of this factor can be determined by testing these ei or eii values.

2.2.2. *Multiple categories*

More categories can be considered, such as three ethnicities and gender. The Zi can be defined as Table 2.

Ethnicities	Gender	Z1	Z2	Z3
A	men	0	0	0
A	women	0	0	1
B	men	1	0	0
B	women	1	0	1
C	men	0	1	0
C	women	0	1	1

Table 2. Example of multiple categories.

The equation to determine the significance of ethnicity and gender is:

$$BF = b_0 + f_1 Z_1 + f_2 Z_2 + f_3 Z_3 + b_1 BMI + g_1 Z_1 BMI + g_2 Z_2 BMI + g_3 Z_3 BMI$$

$$+b_{11} BMI^2 + f_{11} Z_1 BMI^2 + f_{22} Z_2 BMI^2 + f_{33} Z_3 BMI^2 + C_1 age + g_1 Z_1 age$$

$$+ h_2 Z_2 age + h_3 Z_3 age + c_{11} age^2 + h_{11} Z_1 age^2 + h_{22} Z_2 age^2 + h_{33} Z_3 age^2$$

$$+ C_{12} age * BMI + k_1 Z_1 age * BMI + k_2 Z_2 age * BMI + k_3 Z_3 age * BMI \quad (6)$$

2.3. Statistical analysis

This statistical analysis uses sigma plot V.14.0 (SPSS Ing., Chicago, IL, USA).To evaluate a variable's significance, the t-value of its parameter is tested. The hypothesis is.

H0: $b_i = 0$

H1: $b_i \neq 0$

The t-value of b_i is:

$$t = \frac{b_i}{se(bi)} \quad (7)$$

where b_i is the variable's parameter value and se(bi) is the standard error of bi. The coefficient of determination, R2, and the estimation value of standard errors, s, are used as criteria to compare the models' fitting agreement.

The Variance Inflation Factor (VIF) is a statistic that detects the multicollinearity between these independent variables.

2.4. The PRESS statistics

The statistics PRESS is used to evaluate the models' predictive ability[35].

A statistic, the prediction sum of squares (PRESS), is adopted to evaluate the regression model's prediction ability. Considering a data set of size n, PRESS is calculated by omitting each observation individually. Then, the remaining n - 1 observations are used to estimate the coefficients for a candidate model. This regression equation is used to predict the value of the omitted response value (denoted by y^I,-I). We then calculate the ith PRESS residual as the difference yi- y^I,-I. Then, the calculation for PRESS is given by PRESS=∑(yi- y^I,-I)2. The observation yi was not used for fit and assessment, so the ith PRESS residual could serve as a criterion for evaluating prediction ability.

The smaller the PRESS value, the better the model's predictive ability

2.5. BF data source

The data of this study are provided by Dr. Steven B. Heymsfield[39]. All data are from NHANES (National Health and Nutrition Examination Survey) 1999-2006. There are eight data sets in this study. The four ethnicities are White, Black, Mexican, and others. The other group includes Asians, Puerto Ricans, and other countries. However, the number of others is limited. Each ethnicity has two sub-sets of data: men and women. The descriptive statistics of the subjects in different groups are listed in Table 3. The BF was measured using the dual-energy X-ray absorptiometry (DXA) technique.

3. Results

3.1. Descriptive statistics

Table 3 shows the general and anthropometric characteristics of the sample. For the four ethnicities, women had higher mean BF values than men. Mexican women had the highest mean BF values of the other three ethnicities. Similar results were found for the mean BF values.

Sample & variables	Black men	Black women	White man	White women
No. of Subjects	2,286	2,284	2,320	2,366
Age (years)	29.54±20.48	32.53±20.84	43.26±24.13	44.52±23.04
Weight(kg)	73.75±26.25	73.64±24.38	79.50±22.58	70.11±19.06
Stature(cm)	169.80±14.20	160.41±9.23	172.75±12.32	161.58±8.03
BMI(kg/m^2)	24.97±6.88	28.25±8.34	26.21±35.08	26.70±6.66
Waist circumference (cm)	85.26±19.56	90.25±19.45	94.77±17.74	90.55±16.22
BF (%)	24.26±7.85	37.16±8.25	27.85±6.74	38.61±7.44
BF mass(g)	19.26±12.32	29.01±15.08	23.17±10.83	38.16±13.30
Lean soft tissue(g)	52.81±15.88	42.52±10.08	54.48±13.25	40.02±7.74
Total mass(g)	74.80±26.76	73.66±24.59	80.13±22.80	70.15±19.32
	Mexican men	**Mexican women**	**Others men**	**Others women**
No. of Subjects	2,218	2,258	1,043	938
Age(years)	28.39±19.91	33.15±20.85	31.96±20.20	34.22±19.98
Weight(kg)	68.8±20.23	65.69±18.80	71.58±22.04	64.77±19.08
Stature(cm)	163.47±13.22	155.58±8.53	166.89±13.21	156.98±9.01
BMI(kg/m^2)	25.20±6.04	26.91±6.78	25.09±5.74	26.01±6.55
Waist circumference (cm)	88.10±17.40	89.53±16.30	88.14±16.88	87.49±16.20
BF (%)	28.13±7.8	38.87±6.93	26.57±6.94	37.37±7.16
BF mass(g)	20.32±10.54	24.46±11.48	19.81±9.76	24.97±11.49
Lean soft tissue(g)	47.27±13.23	37.24±8.09	45.54±13.72	37.34±8.30
Total mass(g)	69.71±22.45	65.54±18.97	72.03±22.18	64.18±19.90

Table 3. Descriptive statistics of subjects in different ethnic groups.

doi.org/10.32388/T73GYR

3.2. Evaluate the adequate BF equations.

3.2.1. Black men

The relationship between BF and BMI for black men is presented in Figure 1. The BMI and age are assumed as the factors of the BF value.

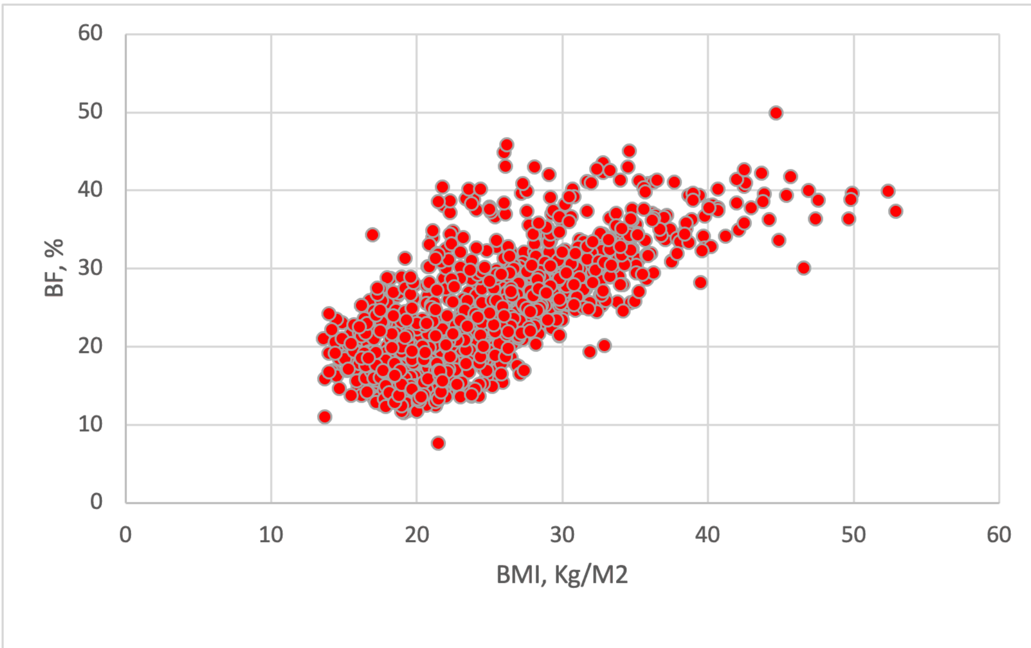


Figure 1. Relationship between BF and BMI for black men.

The results of the multiple regression are as follows:

The members in parentheses below the estimated values of the parameters are the t values for these estimated values.

BF = -0.38 +1.708BMI -0.0126BMI2 – 0.684age + 0.00851age2 +0.0000323BMI*age

(-0.171) (12.301) (-5.47) (-14.29) (17.15) (0.0241) (8)

The t-value for BMI•age was 0.0241. Thus, the term BMI*age had no significant effect on the BF equation.

The adequate equation was as follows:

BF = -0.314 +1.698BMI -0.0125BMI2 – 0.685age + 0.00842age2

(-0.176) (12.320) (-5.328) (-17.792) (14.393)

doi.org/10.32388/T73GYR

$R^2=0.791$, $s=4.675$, PRESS=50,176 (9)

The Kolmogorov-Smirnov test indicated the data distribution of BF was normal. The residual plots of equation (9) are shown in Figure 2a. The uniform distribution of the residuals indicated constant variance.

For the logarithmic BMI value, BMI served as the dependent variable, and the results of the regression analysis are as follows:

$$BF = 37.375 - 28.211 LnBMI + 8.383(LnBMI)^2 - 0.662 age + 0.00821 age^2 \quad (10)$$

$R^2=0.613$, $s=4.691$, PRESS=50,555

The normal test of this equation failed. Figure 2 b shows the residual plot of the BF vs lnBMI equation. The fixed pattern of this residual distribution indicates that the variance of this model is not constant.

When comparing the BF vs. BMI equation, the model had larger s and PRESS values and smaller R^2 values. That is, the fitting agreement and predictive ability were not better than those of the previous BMI model.

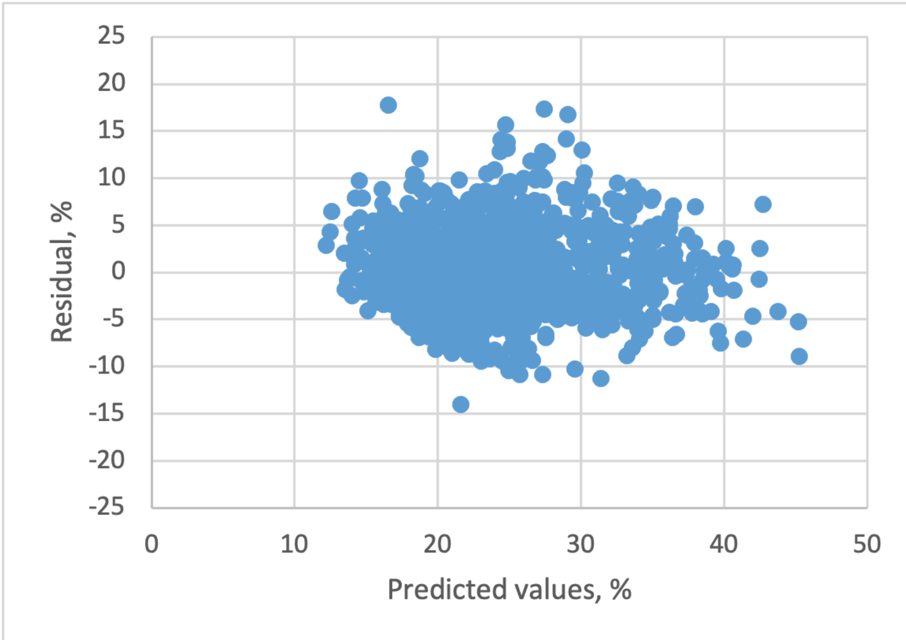
As the inverse BMI served as the new variable, the results of regression analysis are as follows:

$$BF = 85.430 - 1869.322 (1/BMI) + 14321.597((1/BMI)^2 - 0.653 age + 0.00805 age^2 \quad (11)$$

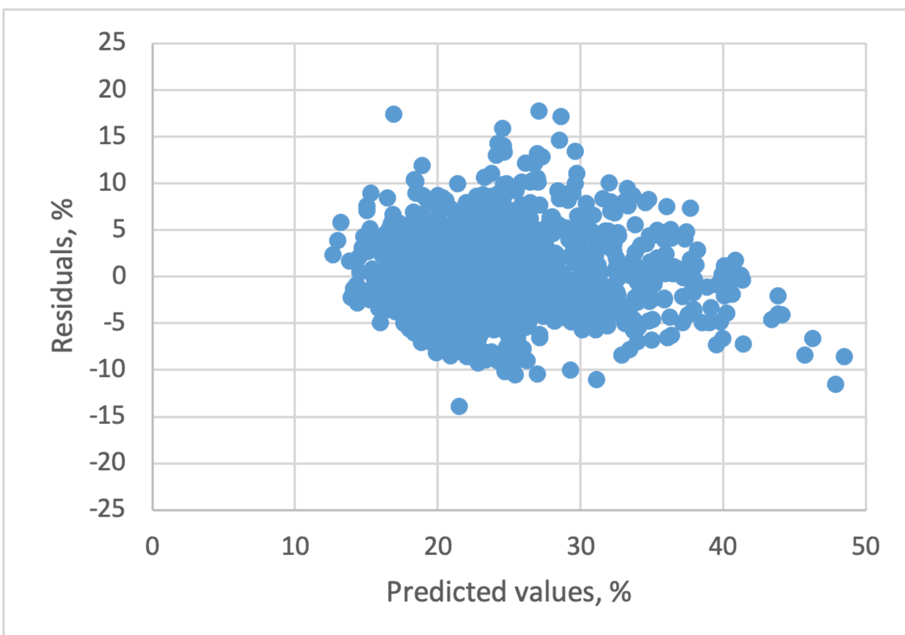
$R^2=0.637$, $s=4.605$, PRESS=49,686

The residual plot of the BF vs. 1/BMI equation is shown in Figure 2c. A fixed pattern could be found, but the variance was not uniform. The normal test failed. At the higher predicted values, the errors were positive. This is a typical overestimation error. Comparing these three equations, the 1/BMI equation had the smallest values of s and PRESS and the most considerable R^2 value. The quantitative criteria of the fitting agreement and perceived performance were close between the BMI and 1/BMI equation. However, only the BMI equation could pass the normal test and homogenous variance tests. It was recognized as an adequate model.

doi.org/10.32388/T73GYR

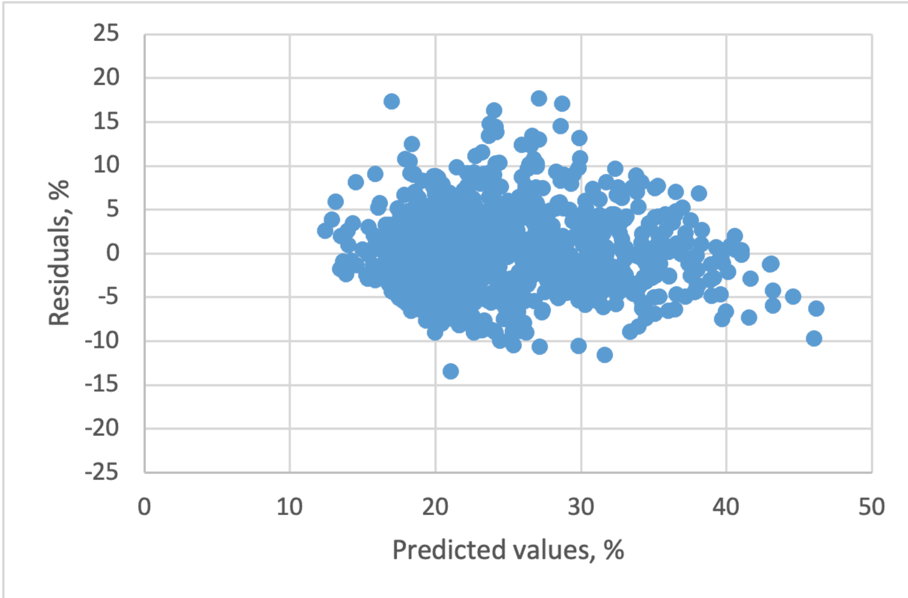


(a). BMI equation



(b). Ln(BMI) equation

doi.org/10.32388/T73GYR



(c). 1/BMI equation

Figure 2. Residual plots of the BF equation for black men.

3.2.2. Black women

The data distribution between BF and BMI for black women is shown in Figure 3.

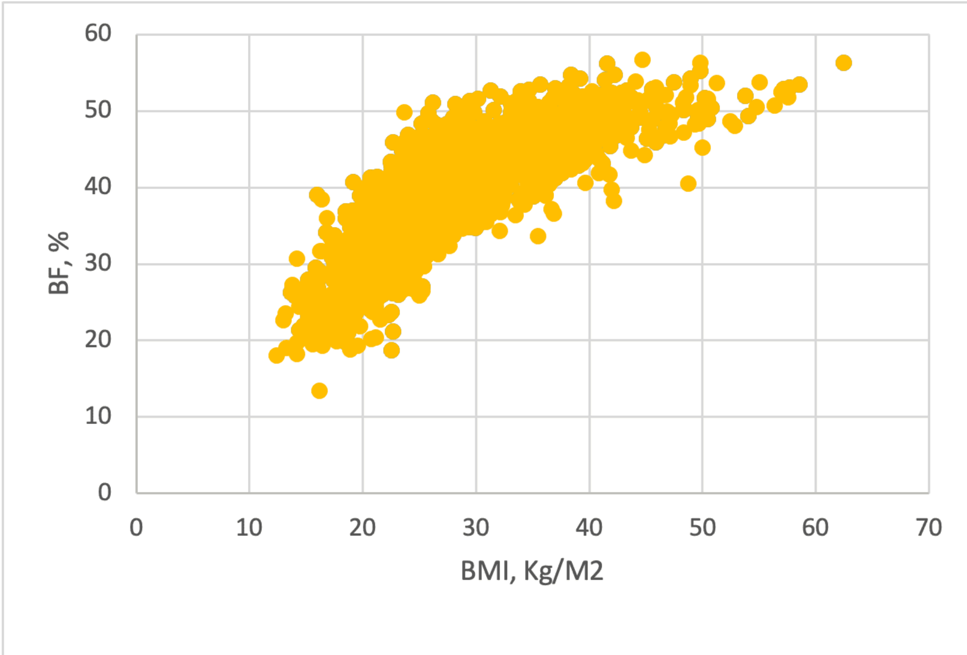


Figure 3. Relationship between BF and BMI for black women.

The regression analysis of BF and its influencing factors are as follows:

BF = -4.371 +2.0878BMI -0.0206BMI2 + 0.0364age + 0.000439age2 -0.0000245BMI*age

(-4.201) (30.858) (-19.462) (1.622) (2.313) (-0.434) (12)

The t-value of the estimated value of BMI*age was -0.434. It did not have a significant effect on the BF. The adequate model is:

BF = -4.207 +2.083BMI -0.0207BMI2 +0.0306age + 0.000432age2

(-4.348) (31.164) (-19.922) (1.698) (12.285)

R2=0.769, s=3.697, PRESS=31,944 (13)

The normal test passed. The residual plots are shown in Figure 4a. The uniform distribution revealed the constant variance.

By transforming the logarithmic BMI value, the lnBMI equation is:

BF = -115.846- 69.069LnBMI -6.843(LnBMI)2 + 0.0159age + 0.00564age2 (14)

R2=0767.613, s =3.708, PRESS=32,362

The equation passed the normal test. The residual plots are shown in Figure 4b. A fixed pattern was found. As the predicted value was below 25%, all residuals were positive and showed over-estimation

doi.org/10.32388/T73GYR

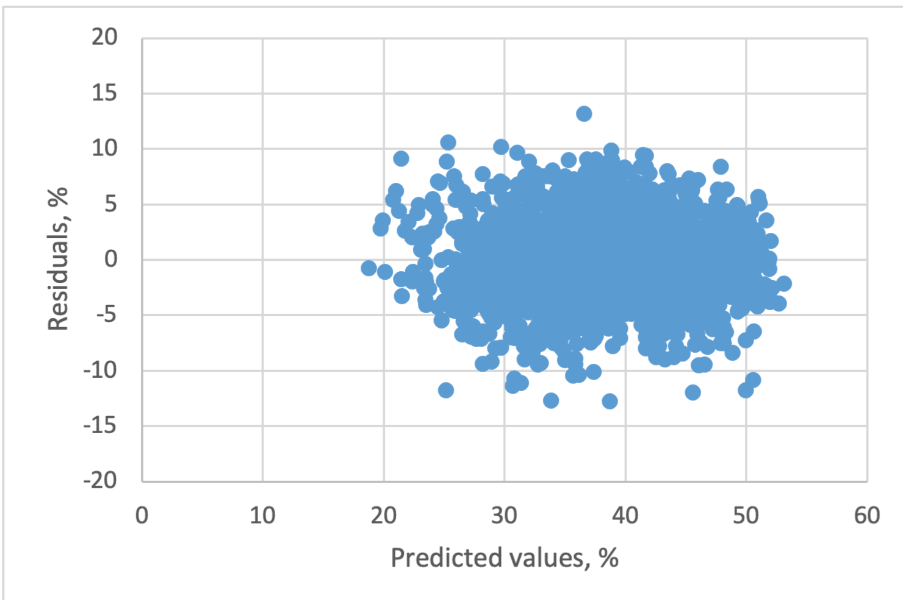
trending. This equation could not be recognized as an adequate model.

The transform form of 1/BMI was used as an independent variable. The regression equation is as follows:

BF = -68.128 – 1011.3 (1/BMI) + 4742.125((1/BMI)2 +0.0232age + 0.000504age2 (15)

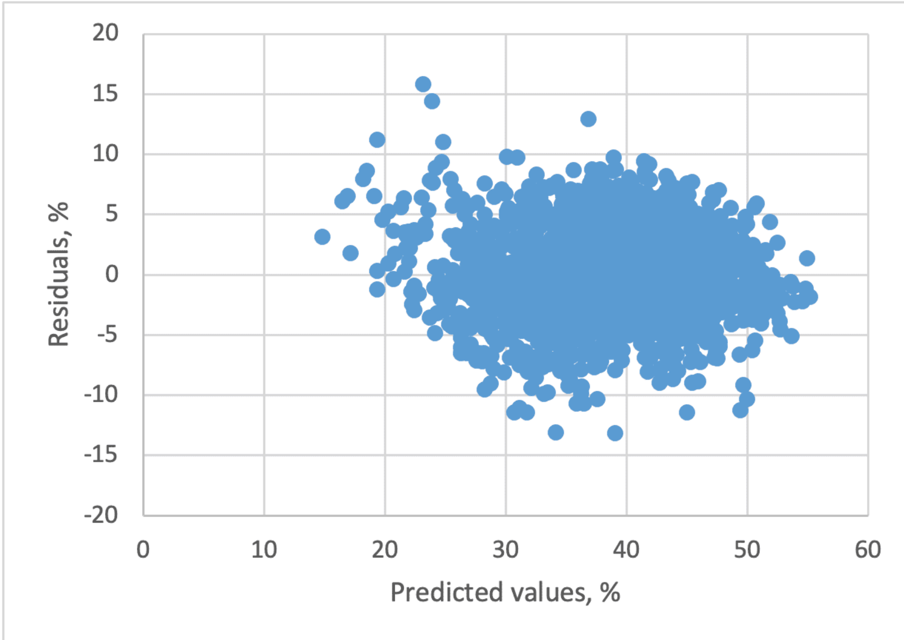
R2=0.769, s=3.697, PRESS=31,843

The normal test passed. The residual plots are presented in Figure 4c. A fixed pattern is found at lower predictive values. Although the R2, s, and PRESS statistics are a little better than that of the BMI equation, it is not deemed a suitable model.

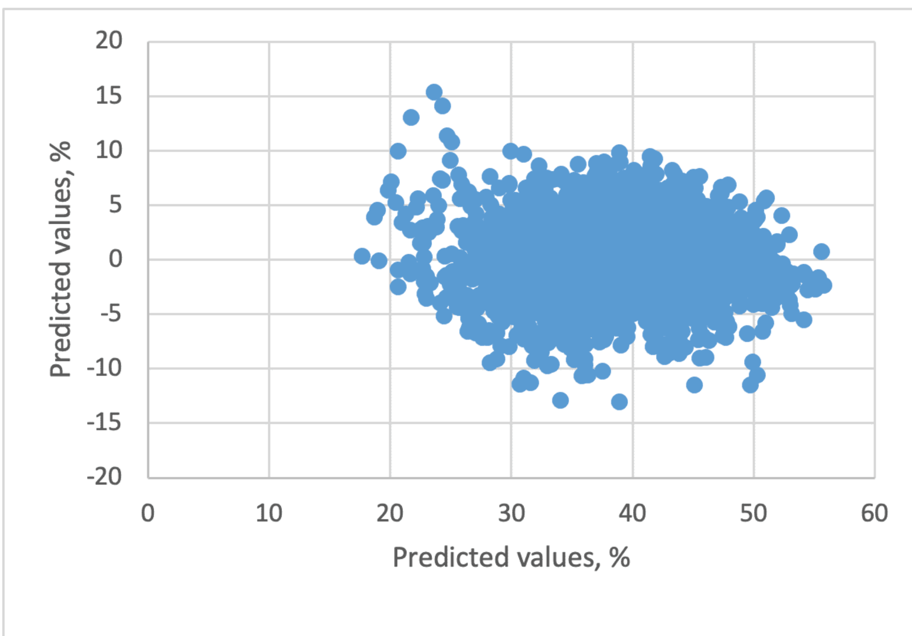


(a). BMI equation

doi.org/10.32388/T73GYR



(b). Ln(BMI) equation



(c). 1/BMI equation

Figure 4. Residual plots of BF equation for black women.

doi.org/10.32388/T73GYR

3.2.3. White men and women

The relationship between BF and BMI for white men and women is presented in Figure 5. At the same BMI value, the BF of women is higher than men.

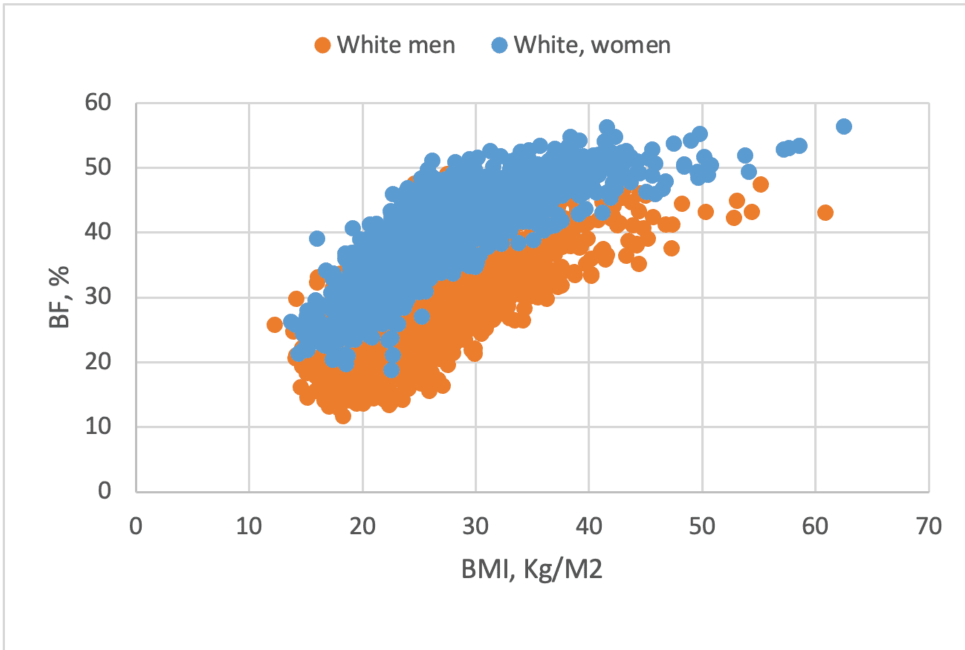


Figure 5. Relationship between BF and BMI for white men and women.

The adequate BF equations are as follows:

White men:

BF = -608 +1.680BMI -0.0124BMI2 -0.411age + 0.00467age2 (16)

R2=0.612, s=4.548, PRESS=44,884

White Women:

BF = -5.831 +2.242BMI -0.0228BMI2 +0.0218age + 0.000377age2 (17)

R2=0.766, s=3.501, PRESS=29,147

3.2.4. Mexican men and women

The data distribution between BF and BMI for Mexican men and women is presented in Figure 6. The adequate models are listed as follows:

doi.org/10.32388/T73GYR

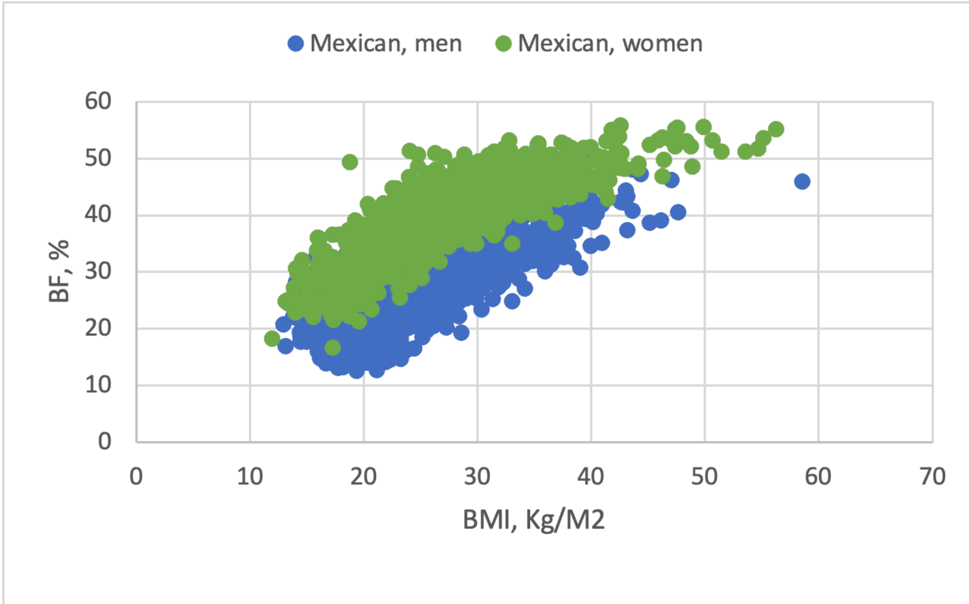


Figure 6. Data distribution between BF and BMI for Mexican men and women.

Mexican men

BF = 1.502 + 1.812BMI - 0.0143BMI2 -0.661age + 0.00758age2 (18)

R2=0.591, s=4.572, PRESS=45,028

BF = 21.013- 16.220LnBMI +6.629(LnBMI)2 - 0.647age + 0.00745age2 (19)

R2=0.588, s=4.589, PRESS=45,405

Mexican women

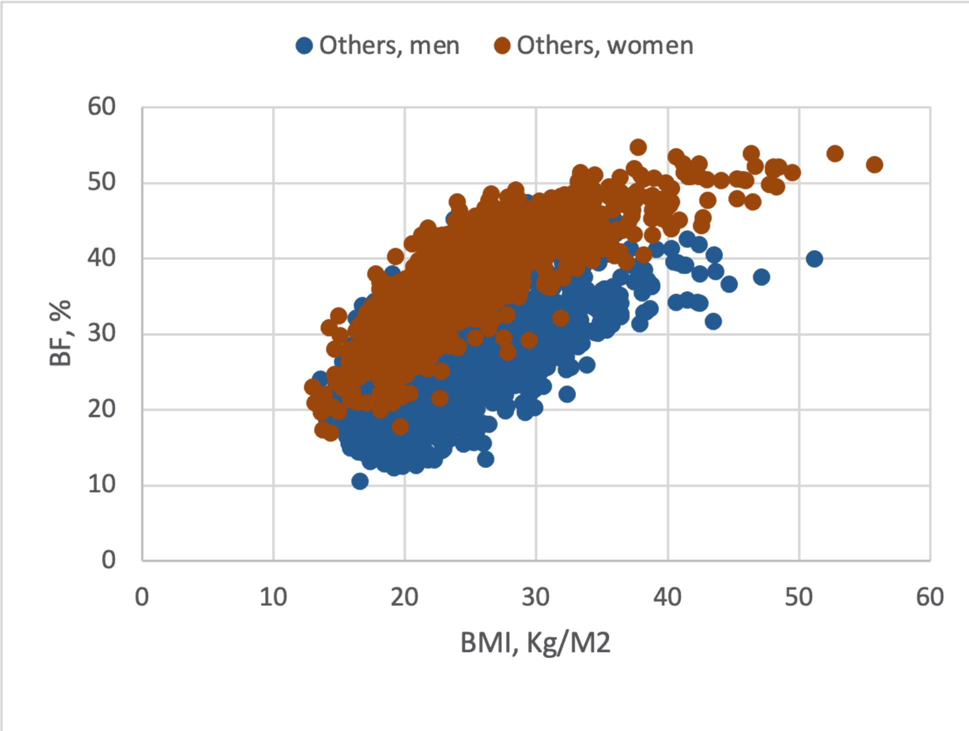
BF = 1.322 + 1.941BMI - 0.0191BMI2 -0.0387age + 0.00097age2 (20)

R2=0.74, s=3.477, PRESS=27,397

3.2.5. Others

The data distribution between BF and BMI for other men and women is presented in Figure 7. The adequate models are:

doi.org/10.32388/T73GYR



Men

BF = 3.055 + 1.502BMI - 0.0110BMI2 -0.478age + 0.00585age2 (21)

R^2=0.489, s=4.957, PRESS=25,733

BF = 30.091- 21.176LnBMI +6.902(LnBMI)2 -0.469age + 0.00577age2 (22)

R^2=0.487, s=4.963, PRESS=25,801

Women

BF = -1.848 + 2.089BMI - 0.0212BMI2 -0.02241age + 0.00758age2 (23)

R^2=0.722, s=3.715, PRESS=13,004

All regression analysis results are listed in Table 4. From the above discussion, the BMI equation could express the relationship between BF and two influence factors (BMI and age) for gender and four ethnicities. The lnBMI equation can be used for Mexican men and other men. However, its fitting agreement and prediction performance were not better than the BMI equations. The BMI equation was then used for the categorical test.

doi.org/10.32388/T73GYR

Criteria Numbers		Black		White		Mexican		Other	
		Men (2286)	**Women (2284)**	**Men (2,320)**	**Wome (2,366)**	**Men (2,218)**	**Women (2,258)**	**Men (1,043)**	**Women (938)**
BMI	R2	0.626	0.798	0.622	0.766	0.591	0.740	0.489	0.722
	s	4.675	3.731	4.492	3.501	4.572	3.477	4.957	3.715
	PRESS	50,176	31,944	44,884	29,149	45,028	27,397	25,733	13,004
	Normal test	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass
	Residual plots	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass
Ln BMI	R2	0.613	40.796	0.609	0.766	0.588	0.741	0.487	0.724
	s	4.691	3.756	4.567	3.503	4.589	3.429	4.963	3.705
	PRESS	50,555	32,362	46,876	29,152	45,405	27,441	25,801	12,929
	Normal test	Fail	Pass	Fail	Pass	Pass	Pass	Pass	Pass
	Residual plots	Fail	Fail	Fail	Fail	Pass	Fail	Pass	Fail
1/BMI	R2	0.637	0.799	0.622	0.760	0.599	0.742	0.493	0.725
	s	4.665	3.725	4.494	3501	4529	3468	4.933	3.701
	PRESS	49,682	31,843	47,579	29,121	44,187	37,377	25,491	12,926
	Normal test	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass
	Residual plots	Fail	Fail	Fail	Fail	Pass	Fail	Pass	Fail

Table 4. Results of regression analysis of BMI, InBMI, and 1/BMI equations for three ethnicities and gender.

3.3. Categorical test of gender

Black men and women

The BMI equation including the gender effect is:

$BF = -0.314 - 4.844Z + 1.698BMI - 0.0125BMI2 + 0.363BMI*Z - 0.00692BMI2*Z$

(-0.197) (-2.434) (13.765) (-5.944) (2.456) (-2.868)

$- 0.684age + 0.0842age2 + 0.696z*age - 0.00807age2*Z$

(-19.851) (19.447) (16.047) (-15.361) (24)

Where Z is the categorized variable, Z=0 is men, and Z=1 is women. The t-values involving the parameters are significant. The gender factor had a significant effect on the BF and BMI and the age relationship of black men and women.

White men and women

The data distribution between BF and BMI for white men and women is in Figure 5. The BMI equation, including the gender effect, is:

$BF = 2.691 – 8.523Z + 1.363BMI - 0.00804BMI2 + 0.879BMI*Z - 0.0148BMI2*Z$

(-4.02) (5.732) (-5.858)

$- 0.333age + 0.00398age2 + 0.355z*age - 0.00361age2*Z$

(9.824) (-9.902) (25)

Where Z is the categorized variable, Z=0 is men, and Z=1 is women. The t-values involving the Z parameters are not equal to zero. Gender had a significant effect on the BMI equation of white people.

Mexican men and women

The BMI equation involving gender factors for Mexican men and women is:

$BF = -0.513 + 3.302Z + 1.976BMI - 0.0172BMI2 - 0.148BMI*Z - 0.00471BMI2*Z$

(4.168) (-4.913) (-1.783)

$- 0.670age + 0.00767age2 + 0.639z*age - 0.00677age2*Z$[26]

The t-values of all parameters involving Z were significant, except the term BMI2*Z. Gender significantly affected the relationship between BF and BMI and age.

Other men and women

The BMI equation involving gender factors for other men and women is:

$BF = 3.055 – 4.903Z + 1.502BMI - 0.0110BMI2 + 0.587BMI*Z - 0.0102BMI2*Z$

(-7.37) (2.703) (-2.710)

$- 0.478age + 0.00585age2 + 0.455z*age - 0.00509age2*Z$

(9.790) (-9.411)[27]

doi.org/10.32388/T73GYR

The t-values of all parameters involving Z were significant, except for the term Z. Gender had a significant effect on the BF value.

From the result of the above discussion, gender is a significant influencing factor on the BF value for four ethnicities: Black, White, Mexican, and others.

3.4. Categorical test of ethnicity

Figures 8 and 9 present the data distribution of BF and BMI for three ethnicities and genders. The significant difference among ethnicities can be observed using the visual method. The categorical test of these data is shown in Table 5.

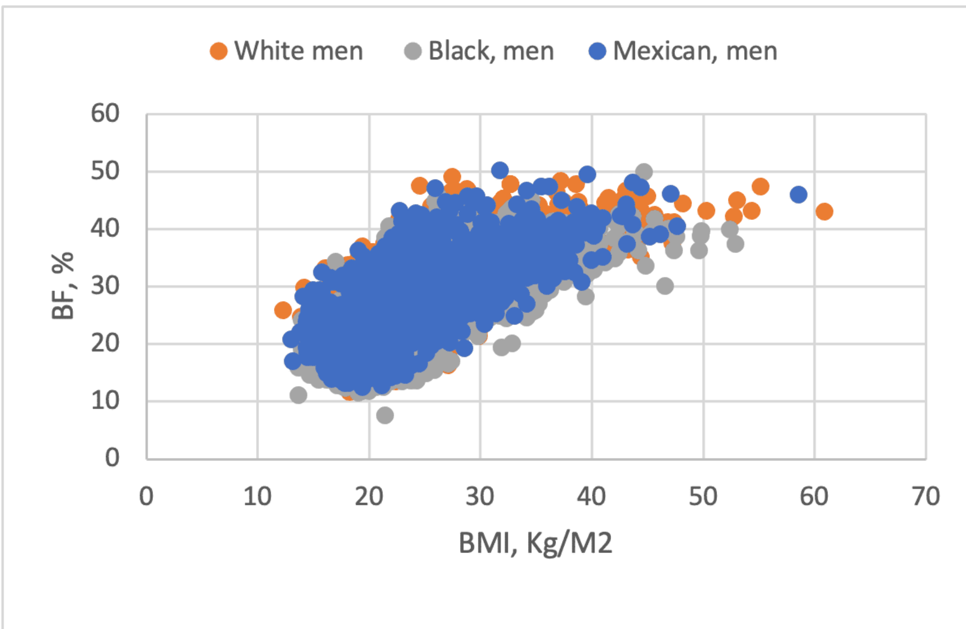


Figure 8. Data distribution of BF and BMI for three ethnicities of men.

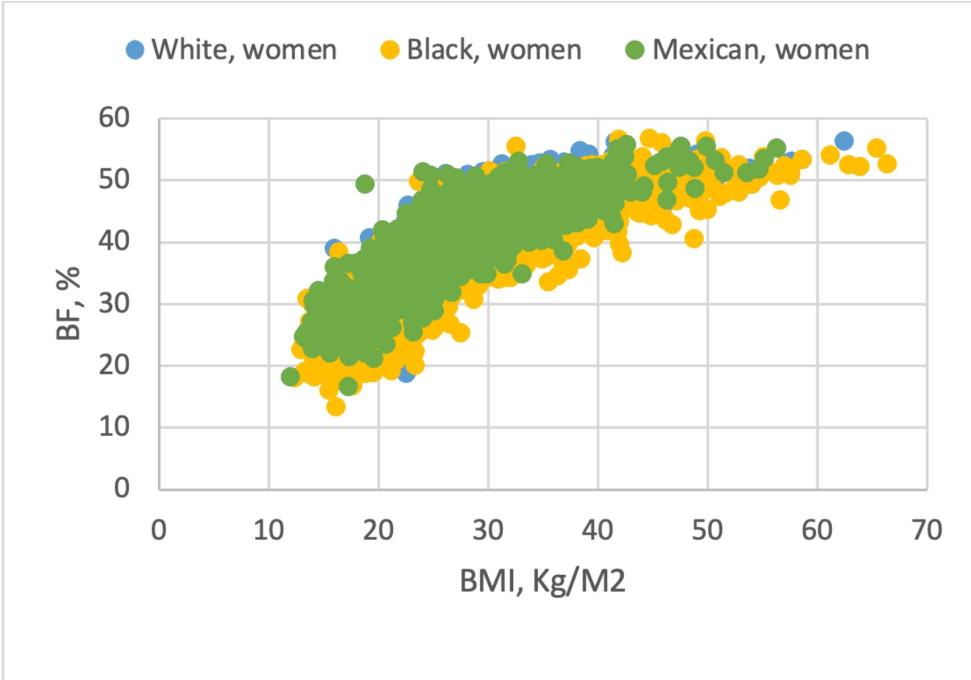


Figure 9. Data distribution of BF and BMI for three ethnicities of women.

Table 5. shows that ethnicity is a significant factor in the BF values.

doi.org/10.32388/T73GYR

1. Men, White and Black

BF=1.064+1.482BMI-0.00997BMI^2+0.114BMI*Z-0.000807BMI2*Z-0.336age+0.0010age2–

(2.117) (-0.708)

0.345age*Z+0.00441age2*Z

(-7.269) (7.895)

Where Z=0 is White men, and Z=1 is Black men.

The categorical factor (Z) significantly affected BMI, age, and age2 variables.

2. Women, White and Black

BF=-5.846+0.974Z+2.266BMI-0.0231BMI^2-0.2278BMI*Z+0.00402BMI2*Z$+0.00989\ age + 0.000450\ age^2$

(-2.811) (2.467)

Where Z=0 is White women, and Z=1 is Black women.

Ethnicity had a significant effect on BMI and BMI2 variables.

However, it did not significantly affect age and age2 valuables.

3. Men, White and Mexican

BF=0.397+3.298Z+1.545BMI-0.0111BMI^2+0.889BMI*Z-0.343age+0.00407age2 – 0.308age*Z+0.00344age2*Z

(3.72) (2.476) (-7.119) (6.951)

Where Z=0 is White men, and Z=1 is White Mexican.

The categorical factor Z significantly affected intercept, BMI, age, and age2 variables.

4. Women, White and Mexican

BF=-5.603+7.070Z+2.267BMI-0.0231BMI^2-0.367BMI*Z+0.00456BMI^2*Z-0.00784age+0.000651age^2

(4.091) (-3.146) (2.293)

Where Z=0 is White women, and Z=1 is Mexican women.

The categorical factor significantly affected intercept, BMI, and BMI2.

However, there was no significant effect on the age and age2 variable.

5. Men, Mexican, and Others.

BF=11.383+8.323Z+1.009BMI-0.0031BMI^2+0.493BMI*Z-0.0119BMI^2*Z

(4.091) (3.146) (2.293)

-0.614age+0.00714age^2-0.138age*Z+0.0013age^2*Z

(-2.765) (-2.178)

doi.org/10.32388/T73GYR

Where Z=0 is Mexican men, and Z=1 is other men. Ethnicity significantly affected intercept, BMI, BMI2, age and age2 .
6. Women, Mexicans and others BF=0.230+-0.615Z+2.001BMI-0.0199BMI^2+0.0302age+0.000864age^2 (-3.955) where Z=0 is Mexican women, and Z=1 is other women. Ethnicity only had a significant effect on intercept.

Table 5. Results of categorical test of BF and BMI and age equations for different ethnicities.

3.5. Categorical test of ethnicity and gender

In this categorical test, a more complex equation (Equation 6) was proposed to consider the effect of ethnicity and gender simultaneously. The factor of ethnicity included three races: Black, Mexican, and White. The sample number of others is limited, so this study did not use the data set. The results of the regression analysis are listed in Table 6.

doi.org/10.32388/T73GYR

BF= 4.538 + 8.277 * z1 + 6.685 * z2 + 1.375 * z3 + 0.818 * BMI - 0.741 * z1*BMI - 0.442 * z2*BMI + 0.691 * z3*BMI + 0.000118 * BMI^2 + 0.0127 * z1*BMI^2 + 0.00737 * z2*BMI^2 - 0.0117 * z3*BMI^2 + 0.0759 * age - 0.0249 * z1*age - 0.109 * z2*age - 0.134 * z3*age - 0.000183 * age^2 + 0.0000113 * z1*age^2 + 0.00119 * z2*age^2 + 0.00140 * z3*age^2				
Coefficient	**Std. Error**	**t**	**P**	
Constant	4.538	0.343	13.242	<0.001
z1	8.277	0.293	28.269	<0.001
z2	6.685	1.607	4.160	<0.001
z3	1.375	1.118	1.229	0.219
BMI	0.818	0.0124	66.187	<0.001
z1*BMI	-0.741	0.0473	-15.677	<0.001
z2*BMI	-0.442	0.119	-3.712	<0.001
z3*BMI	0.691	0.0806	8.573	<0.001
BMI^2	0.000118	0.000390	0.303	0.762
z1*BMI^2	0.0127	0.000937	13.594	<0.001
z2*BMI^2	0.00737	0.00201	3.664	<0.001
z3*BMI^2	-0.0117	0.00133	-8.753	<0.001
age	0.0759	0.00388	19.535	<0.001
z1*age	-0.0249	0.0323	-0.771	0.441
z2*age	-0.109	0.0316	-3. 448	<0.001
z3*age	-0.134	0.0261	-5.136	<0.001
age^2	-0.000183	0.0000736	-2.486	0.013
z1*age^2	0.0000113	0.000347	0.0325	.974
z2*age^2	0.00119	0.000364	3.265	0.001
z3*age^2	0.00140	0.000286	4.887	<0.001

doi.org/10.32388/T73GYR

Table 6. The results of the categorical test of the BMI equation for three ethnicities and gender.

The results of Table 6 indicated that ethnicity and gender both significantly affect the intercept, BMI, BMI2, age, and age2.

The categorical factors, ethnicity, and gender significantly affect the BMI equations. No single form of the BF equation could be proposed to present the relationship between BF, BMI, and age. Each ethnicity of men or women must have its specific BMI equation.

From the above result, all R2 values of female subjects are higher than those of male subjects for the same ethnicity. That is, the variation in BF values could be better explained by BMI and age for women than men. Gender has a significant effect on BF values. At the same BMI value, the BF value of female subjects is higher than that of male subjects.

4. Discussion

In this study, an adequate BF equation is BF = b0 + b1BMI + b11BMI2 + b2age + b22age2. The interaction of BMI and age did not significantly affect the BF value. This model's evaluation includes the normality and constant variance test. All variables were validated using the t-values and p-values of each estimated value of the parameters. After an adequate BF model was established, the effects of sex and ethnicity were tested using a categorical test.

Researchers have studied many forms of BF equations. The simple linear equation validates the narrow BMI range[6][12][39]. When the BMI value is >35 kg/m2, nonlinear or polynomial equations have been proposed[4][11]. Categorical variables such as gender have been incorporated into simple linear equations[7][39]. However, an adequate BF equation was not evaluated in advance. The selection criteria are limited by the classical regression technique, such as R2 and s.

Most of the literature used R2 and s as the criteria to evaluate the fitting agreement of their model. Only some literature mentioned the p-values test to validate the significance of each estimated variable value[8][11][16][19].

Some reports did not test the basic assumptions of the normal distribution and constant variance. The residual plots were only reported by Levitt et al.[4] and Fuster-Parra et al.[18]. In the study of Levitt et

doi.org/10.32388/T73GYR

al.[4], a funnel-type distribution of residuals was found for their nonlinear model, and a fixed pattern of under-estimation was found as BMI > 40 kg/m2. Fuster-Parra et al.[18] proposed a complex form of the BF equation, and log BMI was used as a variable. The residual plots of men and women subjects all revealed a funnel-type distribution. When the BMI > 40 kg/m2, the fixed pattern indicated the predicted values were under-estimated.

The effect of the age variable on the BF value was assumed to be the variable of age in the BF equation[4][7][8]. Only the equation proposed by Meeuwsen et al.[16] considered the power term of age2. However, in our study, the variable of age2 should be incorporated into the BF equation for all ethnicities and both genders.

Meeuwsen et al.[16], Gomze-Ambrosi et al.[17], and Fuster-Parran et al.[18] have introduced complex multiple regression equations. In the study of Merrill et al.[40], four anthropometric measurements—waist circumference, hand thickness, vertical abdominal skinfold, and thigh skinfold—were adopted as the dependent variables. However, the multicollinear tests of these variables were not reported.

The modern regression technique has been used to predict liver volume[41], evaluate the environmental factor on the growth of plantlets in plant tissue culture[42], find the influencing factors on honey properties[43], and evaluate the influencing factors on the dielectric properties of foods[44]. Regression equations with a categorical test could determine which factors significantly affect the BF values. In this study, the significant effects of gender and ethnicity on the BF values are validated with the categorical test of regression analysis. This method can be applied to other data sets of BF, BMI, and other factors to classify these quantitative and qualitative factors.

Faradisa et al.[37] adopted fuzzy logic to establish the relationship between BF and BMI and found that BF and BMI can produce the same categories. However, only classification methods cannot meet the medical requirements for the application of BF values in diagnosis. Xu et al.[38] used a supervised machiner learning technique to developthe prediction equation for BF and BMI and the best model included 18 variables and 19 parameters. The section of the variables in the prediction equations were evaluated by algorithm, a block box for data analysis. There is no academic basis for judging which variables are suitable to be retained in the empirical formula. The best company proposed by the author is derived from the most selected variables but failed to test whether the impact of these variables on BF is significant. In this study, the significant effect for each variable is tested by t-test

doi.org/10.32388/T73GYR

and the predictive performance is evaluated with the PRESS ststistics. These method have its academic basis in Mathematics and Ststistics.

Recently, the BF prediction equation with BMI and other factors still be concerned by researchers.The calculation of the BF values with the BMI and other variables is a simplae way. The prediction ability of the BF empirical equations have been concerned with researchers and clinician. With the study of BF prediction in older adults, Silveria et al[45] suggested to develop specific quations for older adults and the age variables in their application equations did not considering the age and age2 terms. Marin-Jimenez et al.[46] reviewed several BF predicton equations and found the BF calculation equation with anthropometric methods can support a simple and easy indicators and the specific population equations considerded the age and rach should be considered in these equations. Sweatt et al.[47] discussed the strengths and limitations of BMI in the diagnosis of obesity and recommended some anthropomrteies such as waist circumference and waist to heigh ratio to be sonsidered in the BF equations. However, no empirical equations were reported. Takeyama and Fujii[48] recommended that the BF values should considered the effect of height difference of tall, medium, and short. They proposed three BF linear equations and the effect of the heights did not be incoporated in these equations. Taylor et al.[49] report a BF circumference-based equations included the sizes of the abdomen, hip, waist, weist and height. However, they did not reported the ststistics of regression equation. Wu et al.[50] reviewed the advantages and limitations of the application o fthe BMI to evaluate the adult obesity, and confromed that the BMI value is valuable for the primary healthcare screening and limited to served as the predicting of chronic diseases and excess fat assessment. In this study, BF equations have considered the effect BMI, age, gender, and race. The prediction equations developed in this study could be enhance the application of the BMI value with considering other factors.

In our study, humans' BF values are affected by their BMI, BMI2, age, age2, gender, and ethnicity. No single form of the BF equation can be proposed, and cut-off values for BF cannot be selected as fixed values for humans. The modern regresion technique used in this study can be applied to other data sets to establish the adequate BF prediction equation. Then the calculating BF values could served as an index for the primary healthcare screening. With an adequating prediction BF equation, this simple method of the BF calculation with weight, height, and other factors is very helpful for public health management.

doi.org/10.32388/T73GYR

For the future research, the exploring additional anthropometric measures or investigating more complex interactions between factors like physical activity level and body fat prediction, could be provided a broader perspective on the prediction of the BF values.

5. Conclusion

Percent body fat (BF) can serve as an index for human health conditions. A simple predictive equation to calculate the BF value by the influencing factors is very useful. Many empirical equations have been proposed, and inconsistent results have been found. In this study, four data sets, including gender and four ethnicities—white, Black, Mexican, and others—were used to ascertain an adequate BF equation.

The criteria for assessing the appropriate equations used classical regression techniques in previous studies, such as the coefficient of determination, R2, and the estimation value of standard errors, s. In this study, the modern regression technique was used to evaluate the significant effect of each variable with t-test and the prective performance with PRESS ststistics. The effect of other factors such as age, gender, and ethnics was evaluated with categorical testing.

The results of this study indicate the best model of BF uses BMI, BMI2, age and age2 variables. Categorical tests tested the effects of gender and ethnicity on the BF value, and all had a significant effect on BF values. Using a universal BMI equation for all ethnicities or genders is inappropriate. By calculating the BMI value, the cutoff point of BF values needs to consider the difference between gender and ethnicity. This technique in this study provides a reasonable method to establish a BF equation for different age, gender, and ethnicities.

References

1. ^*Levitt DG, Schnider TS. Human physiologically based pharmacokinetic model for propofol. BMC Anesthesiology. 2005;5:4. doi:10.1186/1471-2253-5-4.*
2. ^*Chumlea WC, Guo SS, Kuczmarski RJ, Flegal KM, Johnson CL, Heymsfield SB, Lukaski HC, Friedl K, Hubbard VS. Body composition estimates from NHANES III bioelectrical impedance data. Int J Obes. 2002;26:1596–1609.*
3. ^*Levitt DG, Heymsfield SB, Pierson RN Jr, Shapses SA, Kral JG. Physiological models of body composition and human obesity. Nutr Metab. 2007;4:19-26.*

4. ^a, b, c, d, e, f^ *Levvit DG, Gallagher D, Heymsfield SB. Physiological basis of regression relationship between body mass index (BMI) and body fat fraction. Preedy VR, editor. Handbook of Anthropometry: Physical Measures of Human Form in Health and Disease. Springer; 2012. p. 441-457.*
5. ^ *Ezzati M, Lopez AD, Rodgers A, Vander HS, Murray CJ. Selected major risk factors and global and regional burden of disease. Lancet. 2002;360:1347–1360.*
6. ^a, b, c, d^ *Deurenberg P, Weststrate JA, Seidell JC. Body mass index as a measure of body fatness: Age- and sex-specific prediction formulas. Br J Nutr. 1991;65:105–114.*
7. ^a, b, c, d, e, f^ *Womersley J, Durnin JVGA. A comparison of the skinfold method with extent of 'overweight' and various weight-height relationships in the assessment of obesity. Br J Nutr. 1977;38:271–284.*
8. ^a, b, c, d^ *Gallagher D, Visser M, Sepulveda D, Richard N, Pierson TH, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? Am J Epidemiol. 1996;143:228-239.*
9. ^a, b^ *Deurenberg P, Yap M, van Staveren WA. Body mass index and percent body fat: a meta analysis among different ethnic groups. Int J Obes Relat Metab Disord. 1998;22:1164-1171.*
10. ^a, b, c^ *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;72:694-701.*
11. ^a, b, c, d, e, f^ *Jackson AS, Stanforth PR, Gagnon J, Rankinen T, Leon AS, Rao DC, Skinner JS, Bouchard C, Wilmore JH. The effect of sex, age and race on estimating percentage body fat from body mass index: The heritage family study. Int J Obes Relat Metab Disord. 2002;26:789–796.*
12. ^a, b, c^ *Chang C, Wu C, Chang C, Yao W, Yang Y, Wu J, Lu F. Low body mass index but high percent body fat in Taiwanese subjects: implications of obesity cutoffs. Int J Obes. 2003;27:253–259. doi:10.1038/sj.ijo.802197.*
13. ^a, b, c^ *Kagawa M, Kerr D, Uchida H, Binns CW. Differences in the relationship between BMI and percentage body fat between Japanese and Australian-Caucasian young men. Brit J Nutr. 2006;95:1002–1007.*
14. ^a, b^ *Larson I, Henning B, Lindroos AK, Na¨slund I, Sjöström CD, Sjöström L. Optimized predictions of absolute and relative amounts of body fat from weight, height, other anthropometric predictors, and age 1–3. Am J Clin Nutr. 2006;83:252–259.*
15. ^a, b^ *Rush EC, Goedecke JH, Jennings C, Micklesfield L, Dugas L, Lambert EV, Plank LD. BMI, fat and muscle differences in urban women of five ethnicities from two countries. Int J Obes. 2007;31:1232–1239.*

16. [a], [b], [c], [d], [e]*Meeuwsen S, Horgan GW, Elia M. The relationship between BMI and percent body fat, measured by bioelectrical impedance, in a large adult sample is curvilinear and influenced by age and sex. Clin Nutr. 2010;29:560-566.*

17. [a], [b], [c], [d], [e]*Gomez-Ambrosi J, Silva C, Catalan V, Rodriguez A, Galofre JC, Escalada J, Valenti V, Rotellar F, Romero S, Ramirez B, Salvador J, Fruhbeck G. Clinical usefulness of a new equation for estimating body fat. Diabetes Care. 2012;35:383–388.*

18. [a], [b], [c], [d], [e]*Fuster-Parra P, Bennasar-Veny M, Tauler P, Yañez A, López-González AA, Aguiló A. A comparison between multiple regression models and CUN-BAE equation to predict body fat in adults. PLoS ONE. 2015;10:e0122291. doi:10.1371/journal.pone.0122291.*

19. [a], [b], [c], [d]*Cortés-Castell E, Juste M, Palazón-Bru A, Monge L, Sánchez-Ferrer F, Rizo-Baeza MM. A simple equation to estimate body fat percentage in children with overweightness or obesity: a retrospective study obesity: a retrospective study. Peer J. 2017;5:e3238. doi:10.7717/peerj.3238.*

20. [a], [b], [c]*Itani L, Tannir H, Masri DE, Kreidieh D, Ghoch ME. Development of an easy-to-use prediction equation for body fat percentage based on BMI in overweight and obese Lebanese adults. Diagnostics. 2020;10:728. doi:10.3390/diagnostics10090728.*

21. [a], [b]*Molina-Luque R, Yañez AM, Bennasar-Veny M, Romero-Saldaña M, Molina-Recio G, López-González ÁA. A comparison of equation Córdoba for estimation of body fat (ecore-Bf) with other prediction equations. Int J Environ Res Public Health. 2020;17:7940.*

22. [^]*Duarte MO, Ruelas YF, López-Alcaraz F, del Toro-Equihua M, Sánchez-Ramírez CA. Correlation between percentage of body fat measured by the Slaughter equation and bio impedance analysis technique in Mexican schoolchildren. Nutr Hosp. 2014;29:88-93.*

23. [a], [b]*Ho-Pham LT, Lai TQ, Nguyen MTT, Nguyen TV. Relationship between body mass index and percent body fat in Vietnamese: Implications for the diagnosis of obesity. PLoS ONE. 2015;10:e0127198. doi:10.1371/journal.pone.0127198.*

24. [a], [b]*Akindele MO, Phillips JS, Igumbor EU. The relationship between body fat percentage and body mass index in overweight and obese individuals in an urban African setting. J Public Health Afr. 2016;17:7, 515.*

25. [a], [b]*Jelena J, Baltic ZM, Milica G, Jelena I, Marija B, Milka P, Mirjana L. Relationship between body mass index and body fat percentage among adolescents from Serbian Republic. J Child Obes. 2016;1:5-11.*

26. [a], [b], [c]*Eldeen SKN, Al-Buni R, Al Yami A, Alali H. Relationship between body mass index (BMI) and body fat percentage in a group of Saudi Arabian adults. Ann Public Health Res. 2017;4:1059.*

27. a, b Lorenzo A, Siclari M, Gratteri S, Romano L, Gualtieri P, Marchetti M, Merra G, Colica C. Developing and cross-validation of new equations to estimate fat mass in Italian population. Eur Rev Med Pharmacol Sci. 2019;23:2513–2524.
28. a, b van Haute M, Rondilla E, Vitug JL, Batin KD, Abrugar RE, Quitoriano F, Merced KD, Maaño T, Higa J, Almoro JG, Ternida D, Cabrera JT. Assessment of a proposed BMI formula in predicting body fat percentage among Filipino young adults. Sci Rep. 2020;10:21988.
29. ^ Levitt DG, Heymsfield SB, Pierson RN Jr, Shapses SA, Kral JG. Correction: Physiological models of body composition and human obesity. Nutr Metab. 2009;6:7-11.
30. ^ Chen KK, Wee SL, Pang BWJ, Lau LK, Jabbar KA, Seah WT, Ng TP. Relationship between BMI with percentage body fat and obesity in Singaporean adults–The Yishun Study. BMC Public Health. 2021;21:1-9.
31. ^ Kalavagunta S, Kuppam S, Uma M, Gunda SS, Bhhat A. Correlation between body fat estimated from BMI and bio-impedance analysis among population working in IT sector in Bangalore city –A cross-sectional study. J Evid Based Med Healthc 2021; 8: 1174-1180.
32. a, b Nickerson BS, McLester CN, McLester JR, Kliszczewicz BM. Relative accuracy of anthropometric-based body fat equations in males and females with varying BMI classifications. Clin Nutr ESPEN 2020; 35: 136–140.
33. a, b Jaafar ZA, Kreidieh D, Itani L, Tannir H, Masri DE, Ghoch ME. Cross-validation of prediction equations for estimating the body fat percentage in adults with obesity. Clin Nutr ESPEN 2021; 41: 346-350.
34. a, b de Bortoli AM, de Brito BB, Neves LLV, de Almeida RL, dos Santos L, Barauna VG, Haraguchi FK. Body fat estimated by equations based on anthropometric parameters correlates with bioelectrical impedance in patients undergoing bariatric surgery. J Hum Growth Dev 2022; 32: 185-192.
35. a, b, c Myers RH. Classical and Modern Regression with Applications, 2nd ed.; Duxbury Press: Monterey, CA, USA, 2000.
36. a, b Weisberg S. Applied Linear Regression, 4th ed.; Wiley: New York, NY, USA, 2013; p. 384.
37. a, b Faradisa IS, Muhammad RP, Girindraswari DA. Design body mass index (BMI) and body fat percentage using fuzzy logic. Indonesian J Electronic Electromed Eng Med 2022; 4: 94-106.
38. a, b Xu S, Nianogo RA, Jaga S, Arah OA. Development and validation of a prediction equation for body fat percentage from measured BMI: a supervised machine learning approach. Sci Rep 2023; 13: 8010.
39. a, b, c Heymsfield SB. The data of body fat and body mass index and others. Personal communication. 2020.

40. ^Merrill Z, Chambers A, Cham R. Development and validation of body fat prediction models in America n adults. Obes Sci Pract 2020; 6: 189–195.
41. ^Chen H, Chen C. On the use of modern regression analysis in liver volume prediction equation. J Med I maging Health Inform 2017; 7: 338–349.
42. ^Wang C, Chen C. Use of modern regression analysis in plant tissue culture. Propag Ornam Plants 2017; 17: 83–94.
43. ^Chen C. Relationship between water activity and moisture content in floral honey. Foods 2019; 8: 30.
44. ^Weng YK, Chen J, Cheng CW, Chen C. Use of modern regression analysis in the dielectric properties of fo ods. Foods 2020; 9: 1472.
45. ^Silveira EA, Barbosa LS, Noll M, Pinheiro HA, de Oliveira C. Body fat percentage prediction in older adu lts: Agreement between anthropometric equations and DXA. Clin Nutr 2021; 40: 2091-2099.
46. ^Marin-Jimenez N, Cruz-Leon C, Sanchez-Oliva D, Jimenez-Iglesias J, Caraballo I, Padilla-Moledo C, Cadenas-Sanchez C, Cuenca-Garcia M, Castro-Pinero J. Criterion-related validity of field-based metho ds and equations for body composition estimation in adults: a systematic review. Curr Obes Rep 2022; 1 1: 336-349.
47. ^Sweatt K, Garvey WT, Martins C. Strengths and limitations of BMI in the diagnosis of obesity: What is t he path forward? Curr Obes Rep 2024; 13: 584-595.
48. ^Takeyama Y. Relative composition of body fat percentage to BMI based on height differences. Ann Clin Med Case Rep 2024; 14: 1-6.
49. ^Taylor KM, Castellani MP, Bartlett PM, Oliver TE, McClung HL. Development and cross-validation of a circumference-based predictive equation to estimate body fat in an active population. Obes Sci Pract 202 4; 10: e747.
50. ^Wu Y, Li D, Vermund SH. Advantages and limitations of the body mass index (BMI) to assess adult obes ity. Int J Environ Res Public Health 2024; 21: 757.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.