

1 **Performance evaluation of artificial neural network and multiple linear regression in the**
2 **prediction of body mass index in children**

3
4 Muhammad Asif¹, Ghazi Khan Khosa², Abdullah Mohammad Alomair³, Mohammad Ahmed Alomair³,
5 Muhammad Aslam⁴, Muhammad Arslan⁵, Muhammad Sanaullah⁶, and Justyna Wyszynska⁷

6
7 ¹*Department of Statistics, Govt. Graduate College, Qadir Pur Raan, Multan-Pakistan*

8 ²*Department of pediatric Gastroenterology, The Children Hospital and Institute of Child Health*
9 *Multan- Pakistan.*

10 ³*Department of Quantitative Methods, School of Business, King Faisal University, Al-Ahsa 31982,*
11 *Saudi Arabia.*

12 ⁴*Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan*

13 ⁵*Department of Mathematics and Statistics, Institute of Southern Punjab, Multan, Pakistan*

14 ⁶*Govt Graduate College Muzaffargarh- Pakistan*

15 ⁷*Medical College of Rzeszów University, Institute of Health Sciences, ul. Kopisto 2a, 35-959*
16 *Rzeszów, Poland*

17
18 * Corresponding Author

19 E-mail: asifmalik722@gmail.com (MA)

40 **Performance evaluation of artificial neural network and multiple linear regression in the**
41 **prediction of body mass index in children**
42

43 **Abstract**

44 The body mass index (BMI) provides essential medical information related to body weight for the
45 treatment and prognosis prediction of different diseases. The main goal of the present study was to
46 evaluate the performance of artificial neural network (ANN) and multiple linear regression (MLR)
47 model in the prediction of BMI in children. The data from a total of 5,964 children aged 5 to 12 years
48 were included in study. Age, gender, neck circumference (NC), waist circumference (WC), hip
49 circumference (HpC), and mid upper arm circumference (MUAC) measurements were used to
50 estimate the BMI of children. The ANN and MLR were utilized to predict the BMI. The predictive
51 performance of these methods was also evaluated. Gender-wise average comparison showed that
52 median values of all the anthropometric measurements (except BMI) were significantly higher in
53 boys as compared to girls. For the overall sample, the BMI prediction model was, $\widehat{BMI} = -0.242 -$
54 $0.147 \times \text{Age} - 0.367 \times \text{Gender} + 0.176 \times \text{NC} + 0.041 \times \text{WC} + 0.060 \times \text{HpC} + 0.404 \times \text{MUAC}$. A
55 high R^2 value and lower RMSE, MAPE, and MAD indicated that the ANN is the best method for
56 predicting BMI in children. Our results confirm that the BMI of children can be predicted by using
57 ANN and MLR regression methods. However, the ANN method has a higher predictive performance
58 than MLR.

59

60

61

62

63

64

65 **Introduction**

66 Evaluation of nutritional status is necessary for understanding the health of an individual or
67 population. Children are considered a crucial population in regard to detecting their health status.
68 This is especially important during infancy and adolescence, when the amount of water and adipose
69 tissue undergo considerable changes[1-3]. Body composition changes are directly reflected in
70 anthropometric measures, and the body mass index (BMI) is the most renowned and well-known
71 criterion for determining nutritional status in both children and adults [4]. BMI is calculated by
72 dividing weight in kilograms by height in square meters (i.e., kg/m²), and its cut-offs classify people
73 as underweight, normal weight, overweight or obese [4]. Elevated BMI is strongly linked to the risk
74 of developing cardiovascular health diseases (CHD), including diabetes, hypertension, dyslipidemia,
75 and certain forms of cancer [5].

76 However, the measurements of height or weight could be difficult or even impossible due to
77 the non-availability of accurate portable height scales, weighing machines, and well-trained health
78 workers for collection of measurements. Accordingly, in similar scenarios, the BMI could not be
79 measured. Therefore, the estimation of the BMI preferably to be a simple, affordable method, to be
80 calculated with minimal equipment.

81 To date, various indirect methods have been developed to estimate BMI by measuring
82 different body segments. Several studies [6-9] found that BMI correlates with neck circumference
83 (NC), waist circumference (WC), hip circumference (HpC), and mid-upper arm circumference
84 (MUAC), and Marshall et al. [10] found that measurement of WC, NC and MUAC can be used to
85 estimate BMI accurately. Another study with Pakistani type 2 diabetes patients presented the
86 prediction model for BMI based on WC and HpC measurements using multiple linear regression
87 (MLR) analysis for the BMI prediction [11].

88 Over the last few years, several machine learning (ML) methods have also been applied to
89 predict BMI or obesity efficiently. For example, a more recent study by Lee et al. [12] applied linear
90 regression and different data mining algorithms (i.e., random forest and artificial neural network

91 (ANN) for the prediction of a newborn's BMI based on ultrasound measures and maternal delivery
92 information. Another study used k-nearest neighbor (KNN), classification and regression tree
93 (CART), support vector machine (SVM), and multi-layer perceptron (MLP) algorithms, etc., to
94 predict the BMI based on psychological variables [13]. The BMI of 1568 subjects aged 20-60 years
95 was also predicted using three different data mining algorithms from voice signals information [14].
96 Sancar and Tabrizi [15] used an adaptive neuro fuzzy inference system (ANFIS) to calculate BMI
97 based on metabolic risk factors. The existing literature has shown that no study considered age,
98 gender, and different anthropometric measures as input variables in order to predict BMI in children,
99 and performance of two different prediction methods i.e., ANN and MLR were rarely evaluated. This
100 considerable research gap led to the conduction of this research study. The main goal of the study
101 was to evaluate the performance of artificial neural network (ANN) and multiple linear regression
102 (MLR) model in the prediction of BMI in children. The performance of ANN with MLR were
103 evaluated by estimating the values of co-efficient of determination (R^2), root mean square error
104 (RMSE), mean absolute deviation (MAD) and mean absolute percentage error (MAPE).

105

106 **Materials and methods**

107 A cross-sectional dataset collected from March to June 2016 multi-ethnic anthropometric
108 survey (MEAS) was used in the present study. The data is also publicly available on the Mendeley
109 website at <https://data.mendeley.com/datasets/sxgymx5xjm/1>. A detailed description of the study
110 design, sampling methodology, and inclusion/exclusion criteria of the subjects in this survey has
111 been described elsewhere [6, 16-18]. Briefly, in the MEAS, a total of 10,782 children and
112 adolescents (aged 2-19 years) were recruited and the dataset of school-going children and adolescents
113 (n= 9,929) aged 5 to 19 years was collected from 68 public and private schools. While data for
114 subjects under the age of five were collected in public places such as markets, shopping malls and
115 parks etc. This study only included 5,964 children aged 5 to 12 years. The raw dataset of different

116 anthropometric measurements, i.e., body weight, height, NC, WC, HpC and MUAC, were taken in a
117 comfortable standing position under standard procedure. The complete measurement protocols have
118 been discussed in the previously published studies [6, 16-18]. From the height and weight
119 measurements, BMI was calculated [BMI= weight (kg.) ÷ height (meters)²]. The dependent variable
120 included in the study was the BMI of children while, age, gender of children and measurement of
121 NC, WC, HpC and MUAC, were taken as independent variables.

122 The authors assert that the complete study was conducted in accordance with the ethical
123 standards. The project was approved by the Institutional Ethics Research Board of Bahauddin
124 Zakariya University, Multan under the registration number IRB# Stat-271/2017). Verbal informed
125 consent was obtained from all participants and their parents. Verbal consent was witnessed and
126 formally recorded. Researchers recorded on a form created specifically for documentation of verbal
127 consent the name of the participant/parent who gave the verbal consent. The authors had access to
128 information that could identify individual participants during or after data collection

129

130 Statistical analysis

131 The entire statistical analysis was performed using SPSS ver. 23 (SPSS for Windows,
132 Chicago, IL, USA). The normality of continuous variables (age, BMI, NC, WC, HpC and MUAC)
133 was tested by using the Kolmogorov-Smirnoff test. The significance level of $p < 0.05$ was considered
134 and the results were expressed as median [interquartile range (IQR) = Q_1 - Q_3]. The Mann-Whitney U
135 test was used for average comparisons between the groups. The Spearman's rank correlation (r_s) was
136 used to investigate the correlation between BMI and other anthropometric measurements. Since a
137 correlation of less than 0.3 (i.e., $r < 0.30$) is generally described as a weak correlation[19], therefore,
138 the independent variables having a correlation of less than 0.3 with BMI, were excluded from the
139 analysis. The significance level was set at $\alpha = 5\%$ for the whole analysis.

140 For BMI prediction in children, two different methods, the ANN and MLR were applied. An
141 ANN is a computing system consisting of simple inter-connected processing elements called
142 neurons. The input signals (input data) pass through the network of neurons to generate the network
143 response(s). Each neuron (except the input ones) receives information from several neurons through
144 a connection in proportion to their weights, sums them up and modifies the sum through a non-linear
145 transfer function before passing the signal to other neurons [20]. We also present the block diagram
146 of the current study (Fig 1).

147 **Fig 1. The study block diagram**

148 When modeling the neural network, a multi-layer perceptron with an input layer, hidden layer
149 and an output layer was used. Initially, data on 5,964 children were divided into two different parts,
150 i.e., training data and testing data. The training dataset consists of 70% of all the data (i.e., 4181) and
151 the rest of the data (n=1783) is used as a testing phase. The input data were normalized before
152 training the model. The network was trained in 5000 epochs for different numbers of neurons in the
153 hidden layer. In each epoch, a training data set was selected randomly to prevent learning the
154 especial order of data. The commonly used back-propagation training algorithm “scaled conjugate
155 gradient (SCG)” was used for the training of the models. An activation transfer function called
156 “hyperbolic-tangent” was used in all cases.

157 It is important to adjust the learning rate and momentum terms during the learning process of
158 the neural networks. High weights may destroy the learning behaviour of neural networks. The
159 learning rate is set at a small value to prevent the selection of high weights. Small learning rates
160 slowdowns the learning process. Following Heydari et al. [21], the learning rate and momentum were
161 set at 0.1 and 0.7. The input layer consists of 6 neurons corresponding to independent measures (age,
162 gender, NC, WC, HpC and MUAC). Such measures were used to predict BMI in children. In the
163 hidden layer, different numbers of neurons were used for the optimal selection of network
164 architecture and to prevent overtraining.

165 The MLR analysis was also performed to predict BMI. A linear regression model that involves more
166 than one predictor (regressor) variable is called an MLR model. In an MLR, the relationship between
167 dependent (regressand) and more than one independent variable (s) is expressed by a linear
168 regression equation. An MLR equation with k regressors, is given as under:

$$169 \quad Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i; (i = 1, 2, \dots, n), \quad (1)$$

170 where

171 Y_i = i th observation of the dependent variable,

172 X_{ij} = i th observation of the j th regressor ($j = 1, 2, \dots, k$),

173 β_0 = Y -intercept (the constant term),

174 β_j = regression coefficient corresponding to the j th regressor,

175 ε_i = the error term, assumed to be normal with zero mean and constant variance.

176 Referring to the MLR equation (1), in our study, the MLR equation would be

$$177 \quad BMI_i = \beta_0 + \beta_1 \times Age_i + \beta_2 \times Gender_i + \beta_3 \times NC_i + \beta_4 \times WC_i + \beta_5 \times HpC_i + \beta_6 \\ 178 \quad \times MUAC_i + \varepsilon_i; (i = 1, 2, \dots, n). \quad (2)$$

179 The gender of a child is coded as 1 for a boy and 0 for a girl.

180 The multicollinearity among the regressors was also determined using the variance inflation factor
181 (VIF). The VIF between variables was < 5 , suggesting that multicollinearity was not a problem in the
182 models. For the evaluation of models' prediction performance, different criteria, i.e., RMSE, MAPE,
183 MAD and R^2 were used in the literature [12, 13, 15]. For this study, the model prediction
184 performance was also based on all the latter stated criteria. The model with the highest R^2 , lowest
185 RMSE, MAPE and MAD was chosen as the final predictive model.

186

187

188

189

190 Results

191 The study included 5,964 children (boys = 2865; 48.0% and girls = 3099; 52.0%) with a
192 median age of 9.0 (IQR: 7.0-11.0). The descriptive statistics of anthropometric characteristics and
193 Spearman's correlation coefficients between BMI and NC, WC, HpC, and MUAC were listed (Table
194 1).

195 A sex-based average comparison revealed that the median values of NC, WC, HpC and
196 MUAC were significantly higher in boys than in girls. While the median BMI was not significantly
197 different among the children of both sexes. In the study sample, significant positive correlations were
198 observed between BMI and MUAC ($r = 0.63$), followed by HpC ($r = 0.56$), NC ($r = 0.56$) and WC (r
199 $= 0.51$).

200 In order to predict the BMI values, the proposed MLR model (3) based on explanatory
201 variables i.e., NC, WC, HpC, MUAC, age and gender (boys =1, girls=0) was used (Table 2).

$$202 \quad \widehat{BMI} = -0.242 - 0.147 \times \text{Age} - 0.367 \times \text{Gender} + 0.176 \times \text{NC} + 0.041 \times \text{WC} + 0.060 \times \\ 203 \quad \text{HpC} + 0.404 \times \text{MUAC}$$

$$204 \quad (3)$$

205 For instance, the MLR model predicts a BMI value of 13.08 of a boy having age (5), NC (21.59),
206 WC (50.80), HpC (48.26) and MUAC (13.97). Similarly, a BMI value of 14.97 was predicted for a
207 girl having age (12), NC (26.67), WC (50.80), HpC (57.15) and MUAC (16.76). Moreover, an ANN
208 model predicts the BMI values for the same boys and girls to be 13.13 and 15.52, respectively (see
209 Table 3).

210 The value of co-efficient of determination (R^2) using MLR analysis revealed that about
211 48.0% variation in BMI is explained due to the predictor variables including the age and sex. While
212 R^2 using the ANN algorithm exhibited that more (53.4%) variation in BMI is explained due to the
213 predictor variables. The RMSE, MAPE, and MAD values were also found to be lower in the ANN as

214 compared to the MLR model. Bases on these findings, we can conclude that ANN outperforms MLR
215 in predicting BMI in children (Table 4).

216 **Discussion**

217 Obesity in children has now evolved into a severe public health issue, and its incidence has
218 grown rapidly in recent years across the world [4]. Researchers used BMI as internationally accepted
219 measure for defining overweight and obesity in both children and adults [22, 23]. Different studies in
220 recent years have also utilized some other anthropometric measurements, i.e., WC, MUAC and NC
221 for obesity screening purposes in children [7-9]. Because these measurements had a good correlation
222 with BMI ($r = \sim 0.60$ to ~ 0.85). Based on receiver operating characteristics (ROC) curves analysis,
223 the diagnostic ability of WC, MUAC and NC to detect children with overweight and obesity was
224 very high i.e., areas under the curve (AUC) values between 75.0 % to 97.0 % [7-9]. Therefore, these
225 measurements can be used to predict BMI in children. To the author's knowledge, this is the first
226 study that explores whether BMI in children can be predicted from anthropometric measurements by
227 using MLR and ANN algorithms.

228 In this study, the neural network was designed using the information of 5,964 children living
229 in different cities of Pakistan, including six input variables (age, gender, NC, WC, HpC and MUAC)
230 and BMI as the output variable. A high R^2 and lower RMSE, MAPE values indicated that ANN is the
231 best method for predicting BMI in children than MLR. These findings are consistent with the earlier
232 reports on the topic [24, 25]. In a study of 321 adult individuals in Iran, an RMSE (0.94) and R^2
233 (0.890) using ANN were much better than MLR (RMSE= 1.31 and $R^2= 0.882$) [24]. Another Iranian
234 study estimated the BMI for 470 adult individuals and reported low RMSE values as compared to our
235 results [25]. A high disparity in R^2 and RMSE results may be due to the fact that they predicted the
236 BMI for adult individuals based on different metabolic syndrome components (i.e., WC, SBP, DBP,
237 FG, HDL and TG) and on different environmental and physical activity-related factors. However, our
238 study predicted the BMI for children aged 5-12 years based on age, gender and anthropometric-

239 related information as input variables. Some studies also employed the ANN method for obesity
240 prediction e.g., an Iranian study with 414 adults found that ANN with an accuracy of 81.2% is a
241 more efficient method for obesity prediction than logistic regression (accuracy = 80.2%) [21].
242 Another study compared the performance of ANN and logistic regression methods for obesity
243 prediction among 82 individuals, and found that ANN performed better than logistic regression [26].
244 Using MLR analysis, we also found that NC, WC, HpC and MUAC are significant predictors for
245 predicting BMI. These results are consistent with an earlier study by Marshall et al. [10], which
246 reported that measurements of WC, NC and AC can be used to accurately estimate BMI and another
247 study with 24,485 Pakistani type 2 diabetes patients aged 20 years and above also offered a BMI
248 prediction model based on WC and HpC measurements [11].

249 The major strength of the study is that we predicted the BMI in children based on NC, WC,
250 HpC and MUAC whose measurements are simple, quick, and just require a non-stretchable plastic
251 tape. Our recommendation is to include the metabolic risk-related variables like SBP, DBP, FG,
252 LDL, HDL and TG that affect BMI. Some studies predicted the BMI by using voice signals
253 information [14], psychological [13], environmental and physical activity-related variables [25], and
254 it would be an important contribution to extend this study based on these new variables. Lastly, the
255 focus of this study was to predict BMI for children and however BMI estimation for adult individuals
256 need to be studied and it will be a good topic for future research. However, the study has limitations.
257 Since our study has a cross-sectional character, we cannot conclude a cause-and-effect relationship.
258 The limitation of this study was a single anthropometric measurements, thus intra-observer
259 variability could not be calculated. However, these measurement was always performed by the same
260 researcher and there was no inter-observer variability between measurements.

261 **Conclusion**

262 The findings of this study imply that both methods, MLR and ANN can be used to predict
263 BMI in children. The use of ANN, a high R^2 and lower RMSE, MAPE and MAD values demonstrate

264 that this method is biologically acceptable and more effective for predicting BMI based on age,
265 gender and four different anthropometric variables. Our methods and results can be used for obesity
266 prediction in Pakistani children as an alternative to the clinical findings and public health research.
267 Further research to overcome the present study's limitations is also required.

268 **Acknowledgments**

269 The author(s) are thankful to Mr. Muhammad Qasim, who reviewed the statistical interpretation and
270
271 made corrections if required.

272 **The author (s) received no specific funding for this work.**

273 **References**

- 274
275 1. Mihalopoulos NL, Holubkov R, Young P, Dai S, Labarthe DR. Expected changes in clinical
276 measures of adiposity during puberty. *J Adolesc Health* 2010; 47:360-66.
- 277
278 2. Talwar I, Sharma K, Kapur S. Growth trends in body, fat, circumferential and physiological
279 traits during adolescents among Rajput females of Theog, Shimla District, India. *Ann Hum*
280 *Biol* 2010; 37:536-53.
- 281
282 3. Wells JC, Fuller NJ, Dewit O, Fewtrell MS, Elia M, Cole TJ. Four component model of body
283 composition in children: density and hydration of fat-free mass and comparison with simpler
284 models. *Am J Clin Nutr* 1999; 69:904-12.
- 285
286 4. WHO Consultation on Obesity (1999: Geneva, Switzerland) & World Health Organization.
287 (2000). Obesity : preventing and managing the global epidemic : report of a WHO
288 consultation. World Health Organization. <https://apps.who.int/iris/handle/10665/42330>
- 289
290 5. Field AE, Coakley EH, Must A, Spadano JL, Laird N, Dietz WH, et al. Impact of overweight
291 on the risk of developing common chronic diseases during a 10-year period. *Arch Int Med*
2001; 161(13):1581-86.

- 290 6. Asif M, Aslam M, Mustafa S, Alfat S. Age-specific Differences and Interrelations between
291 Anthropometric Variables in Pakistani Children aged 2 to 19 years. *Rawal Med J* 2018;
292 43:164-69.
- 293 7. Nafiu OO, Burke C, Lee J, Voepel-Lewis T, Malviya S, Tremper KK. Neck circumference
294 as a screening measure for identifying children with high body mass index. *Pediatrics* 2010;
295 126(2):e306-10.
- 296 8. Mazıcıoğlu MM, Hatipoğlu N, Oztürk A, Çiçek B, Ustünbaş HB, Kurtoğlu S. Waist
297 circumference and mid-upper arm circumference in evaluation of obesity in children aged
298 between 6 and 17 years. *J Clin Res Pediatr Endocrinol* 2010; 2:144-50.
- 299 9. Asif M, Aslam M, Altaf, S. Mid-upper-arm circumference as a screening measure for
300 identifying children with elevated body mass index: a study for Pakistan. *Korea J Pediatr*
301 2018; 61:6-11.
- 302 10. Marshall AR, Haboubi N, Jones S. Body mass index estimation from waist, neck and mid-
303 arm circumference. *Gastrointestinal Nurs* 2011; 9:37-40.
- 304 11. Ghias M, Khawaja KI, Masud F, Atiq S, Khalid Pervaiz M. A new approach for estimation
305 of body mass index using waist and hip circumference in type 2 diabetes patients. *J Ayub*
306 *Med Coll Abbottabad* 2010; 22:111-116.
- 307 12. Lee KS, Kim HY, Lee SJ, Kwon SO, Na S, Hwang HS, et al. Korean Society of Ultrasound
308 in Obstetrics and Gynecology Research Group. Prediction of newborn's body mass index
309 using nationwide multicenter ultrasound data: a machine-learning study. *BMC Pregnancy*
310 *Childbirth* 2021; 21(1):172.
- 311 13. Delnevo G, Mancini G, Rocchetti M, Salomoni P, Trombini E, Andrei F. The Prediction of
312 Body Mass Index from Negative Affectivity through Machine Learning: A Confirmatory
313 Study. *Sensors (Basel)* 2021; 21(7):2361.

- 314 14. Lee BJ, Kim KH, Ku B, Jang JS, Kim JY. Prediction of body mass index status from voice
315 signals based on machine learning for automated medical applications. *Artif Intell Med*
316 2013; 58(1):51-61.
- 317 15. Sancar N, Tabrizi SS. Body mass index estimation by using adaptive neuro fuzzy inference
318 system. *Procedia Comput Sci* 2017; 108C:2501-06.
- 319 16. Asif M, Aslam M, Qasim M, Altaf S, Ismail A, Ali H. A dataset about anthropometric
320 measurements of the Pakistani children and adolescents using a cross-sectional multi-ethnic
321 anthropometric survey. *Data Brief* 2020; 34:106642.
- 322 17. Asif M, Aslam M, Altaf S, Mustafa S. Developing waist circumference, waist-to-height ratio
323 percentile curves for Pakistani children and adolescents aged 2-18 years using Lambda-Mu-
324 Sigma (LMS) method. *J Pediatr Endocrinol Metab* 2020; 33:983-93.
- 325 18. Asif M, Aslam M, Khan S, Altaf S, Ahmad S, Qasim M, Ali H, Wyszynska J. Developing
326 neck circumference growth reference charts for Pakistani children and adolescents using the
327 lambda-mu-sigma and quantile regression method. *Public Health Nutr* 2021; 24(17):5641-49.
- 328 19. Profilldis VA, and Botzoris GN. Modeling of Transport Demand, Analyzing, Calculating,
329 and Forecasting Transport Demand. Ch # 5, 2019, pp. 163-224. DOI:
330 <https://doi.org/10.1016/B978-0-12-811513-8.00005-4>.
- 331 20. Dreyfus G. *Neural networks: methodology and applications*. Springer Verlag, 2005. DOI:
332 10.1007/3-540-28847-3.
- 333 21. Heydari ST, Ayatollahi SM T, Zare N. Comparison of Artificial Neural Networks with
334 Logistic Regression for Detection of Obesity. *J Med Syst* 2012; 36:2449-54.
- 335 22. Ogden CL, Carroll MD, Flegal KM. High body mass index for age among US children and
336 adolescents, 2003-2006. *JAMA* 2008; 299:2401-5.
- 337 23. Lobstein T, Baur L, Uauy R. Obesity in children and young people: a crisis in public health.
338 *Obes Rev* 2004, 5,4-85.

- 339 24. Tabrizi SS, Sancar N. Prediction of Body Mass Index: A comparative study of multiple
 340 linear regression, ANN and ANFIS models. *Procedia Com Sci* 2017; 120:394-401.
- 341 25. Hoseini SH, Soltani A, Pourahmadi-Nakhli M. Application of artificial neural network
 342 (ANN) in estimation of body mass index (BMI) based on the connection between
 343 environmental factors and physical activity. *Int J Artif Intell* 2012; 3:107.
- 344 26. Ergün U. The classification of obesity disease in logistic regression and neural network
 345 methods. *J Med Syst* 2009; 33:67-72.

346

347 Table 1. Descriptive statistics and Spearman's correlation coefficients between BMI and different
 348 anthropometric characteristics of the study subjects.

349

Characteristics	Total (n=5964)	Boys (n=2865)	Girls (n=3099)	Correlation (r _s)
	Median (IQR)	Median (IQR)	Median (IQR)	
BMI (kg/m ²)	15.26 (14.05-16.91)	15.23 (14.09-16.76)	15.26 (13.97-17.05) ^{NS}	1.00
NC (cm)	25.40 (24.13-26.67)	25.45 (24.13-26.67)	25.40 (23.37-26.67) ^S	0.55 ^S
WC (cm)	55.88 (50.80-60.96)	56.39 (52.07-62.23)	55.88 (50.80-60.96) ^S	0.51 ^S
HpC (cm)	60.96 (55.88-67.31)	63.50 (58.42-68.58)	60.96 (55.88-66.04) ^S	0.56 ^S
MUAC (cm)	16.51 (15.24-17.78)	16.54 (15.24-17.79)	16.51 (15.24-17.78) ^S	0.63 ^S

IQR: Interquartile range; BMI: Body mass index; NC: Neck circumference; WC: Waist circumference; HpC: Hip circumference; MUAC: Mid-upper arm circumference. S: Significant < 0.001; NS: Not significant

350

351

352

353

354

355

356

357 Table 2. Estimated regression co-coefficients by using multiple linear regression.

Characteristics	β	t-statistic	SE	95% CI for β	p-value
Age	-0.147	-11.48	0.013	(- 0.17, - 0.12)	<0.001
Gender	-0.367	-7.73	0.047	(- 0.46, - 0.27)	<0.001
NC	0.176	11.05	0.016	(0.14, 0.21)	<0.001
WC	0.041	9.11	0.004	(0.03, 0.05)	<0.001
HpC	0.060	11.14	0.005	(0.05, 0.07)	<0.001
MUAC	0.404	24.38	0.017	(0.37, 0.44)	<0.001
Constant	-0.242	-0.88	0.275	(- 0.78, 0.30)	0.380

358 β : regression co-efficient; NC: neck circumference; WC: waist circumference; HpC: hip circumference.
 359 MUAC: mid-upper arm circumference; SE: standard error

360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382

383
384
385
386
387
388

Table 3. Comparison of observed and estimated BMI by using multiple linear regression and artificial neural network

Age	Gender	NC	WC	HpC	MUAC	Observed BMI	Estimated	Estimated
							BMI using MLR	BMI using ANN
5	0	25.40	45.72	53.34	13.97	16.66	14.21	13.85
5	1	21.59	50.80	48.26	13.97	15.00	13.08	13.13
5	1	18.80	50.80	48.26	12.70	13.46	12.07	12.65
6	1	23.11	43.18	52.07	15.24	15.13	13.63	14.54
6	1	22.35	43.18	50.80	13.21	14.79	12.60	13.60
6	0	24.13	55.88	53.34	15.24	12.93	14.77	14.18
6	0	24.13	45.72	53.34	15.24	14.12	14.35	14.64
7	0	24.13	63.50	58.42	15.24	13.44	15.24	14.86
7	1	25.40	55.88	55.88	15.75	16.39	14.84	14.53
7	1	24.13	53.34	58.42	16.51	14.54	14.97	14.89
8	1	27.94	63.50	73.66	20.83	18.66	18.57	19.39
8	1	22.86	58.42	55.88	17.27	13.88	14.96	14.72
9	1	25.15	62.23	62.23	17.78	15.62	15.96	15.62
9	1	22.86	50.80	58.42	15.24	14.35	13.84	14.31
9	1	27.94	55.88	60.96	15.24	15.75	15.09	15.00
10	1	26.67	66.04	68.58	19.05	15.82	17.13	17.02
10	1	24.89	50.80	58.42	15.75	13.82	14.25	14.43
11	1	25.15	58.42	60.96	17.78	17.09	15.44	15.04
11	1	25.91	68.58	71.12	21.59	22.68	18.14	18.39
12	1	29.21	66.04	73.66	21.59	17.67	18.62	18.47

12	0	22.86	49.53	60.96	16.51	13.27	14.38	14.97
12	0	26.67	50.80	57.15	16.76	14.88	14.97	15.52
12	0	27.94	53.34	90.42	18.80	16.15	18.12	18.58

389 Gender (1=boys, 0= girls); NC: Neck circumference; WC: Waist circumference; HpC: Hip circumference; MUAC: mid-
390 upper arm circumference; MLR: Multiple linear regression; ANN: Artificial neural network

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415 Table 4. Predictive performance of multiple linear regression and artificial neural network.

Algorithms	Model Evaluation Criteria			
	R²	RMSE	MAPE	MAD
ANN	0.534	1.7047	0.0821	1.2758
MLR	0.4782	1.7953	0.0810	1.2761

416 ANN: Artificial neural network; MLR: Multiple linear regression; R²: Co-efficient of determination; RMSE: Root mean
417 square error; MAPE: Mean absolute percentage error; MAD: Mean absolute deviation.