UDC 612.3-004.42
DOI: 10.21668/health.risk/2024.1.16.eng

Read online

Research article

IDENTIFYING THE FACTORS RELATED TO BODY FAT PERCENTAGE AMONG VIETNAMESE ADOLESCENTS USING MACHINE LEARNING TECHNIQUES

Nguyen Thi Hong Hanh[1], Le Thi Tuyet[1], Nguyen Thi Trung Thu[1], Do Thi Nhu Trang[1], Duong Thi Anh Dao[1], Le Thi Thuy Dung[2], Dang Xuan Tho[3]

[1]Hanoi National University of Education, 136 Xuan Thuy St., Hanoi, Vietnam
[2]Binh Duong General Hospital, 5 Pham Ngoc Thach, Hiep Thanh, Binh Duong, Vietnam
[3]Academy of Policy and Development, Hoai Duc district, Hanoi, Vietnam

The aim of this study was to investigate the factors influencing Body Fat Percentage (BFP) among Vietnamese adolescents aged 11 to 15 employing machine learning techniques for predictive analysis.

A total of 1,208 adolescents, comprising 598 boys and 610 girls, drawn from nine junior high schools in Vietnam's capital, were enrolled in the study. Body composition measurements were conducted using the HBF 375 (Omron) device by Bioelectrical Impedance Analysis method. The study questionnaire, initially validated by The National Institute of Nutrition, encompassed inquiries related to dietary behaviors, meal frequencies, physical activities, sedentary habits, and nutritional knowledge. A machine learning methodology employing a decision tree algorithm was employed to discern the primary determinants most significantly correlated with BFP.

This study successfully identified six distinct predictor groups associated with BFP among adolescents, leveraging the decision tree model, with Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) values of 4.80 and 3.80, respectively. Among these predictors, frequency of fruit consumption, snacking habits, mode of transportation to school, and screen time (computer and/or cell phone usage) emerged as the most influential factors linked to BFP in adolescents. The combination of these factors and interactions with gender and pubertal status can BFP in Vietnamese adolescents.

This research sheds light on the complex and diverse factors impacting BFP in Vietnamese adolescents. This study's results underscore the practical importance of promoting healthy eating and exercise habits among adolescents, offering valuable insights for parents and schools to enhance their childcare strategies.

Keywords: *machine learning, body fat percentage, predictability, influencing factors, eating habits, physical activity,Vietnamese adolescents, the decision tree.*

Body Fat Percentage (BFP) is widely accepted as an accurate and effective measurement of obesity status. Previous evidence has suggested that this index can provide additional insights into risks of cardiovascular health and other metabolic disorders as adipokines secreted by adipose tissue may affect many metabolic functions such as fat distribution, appetite, energy expenditure, insulin sensitivity and secretion, glucose and

Nguyen Thi Hong Hanh – PhD., Lecturer at the Department of Biology (e-mail: hanhnth@hnue.edu.vn; tel.: +84 (902) 80-11-83; ORCID: https://orcid.org/0000-0002-1392-8081).

Le Thi Tuyet – PhD., Lecturer for Human and Animal Physiology at the Department of Biology (e-mail: tuyetlt@hnue.edu.vn; tel.: +84 (968) 79-55-55; ORCID: https://orcid.org/0000-0002-3308-5886).

Nguyen Thi Trung Thu – PhD., Lecturer at the Department of Biology (e-mail: trungthu@hnue.edu.vn; tel.: +84 (983) 42-09-85; ORCID: https://orcid.org/0000-0002-9800-2287).

Do Thi Nhu Trang – M.S., Assistant Lecturer at the Department of Biology (e-mail: trangdtn@hnue.edu.vn; tel.: +84 (986) 70-38-10; ORCID: https://orcid.org/0009-0009-5461-4554).

Duong Thi Anh Dao – PhD., Associate Professor, Lecturer at the Department of Biology (e-mail: daodangduc@gmail.com; tel.: +84 (886) 81-67-66; ORCID: https://orcid.org/0009-0003-4407-8942).

Le Thi Thuy Dung – PhD., M.D. at the Department of Pediatrics (e-mail: letono2002@gmail.com; tel.: +84 (987) 00-89-14; ORCID: https://orcid.org/0000-0002-8855-7801).

Dang Xuan Tho – PhD., Main Lecturer at the Faculty of Digital Economics (tel.: +84 (912) 62-93-83; ORCID: https://orcid.org/0000-0002-7654-5942).

lipid metabolism [1]. The association between BFP and risk factors related to obesity is a subject of controversy when considering diverse ethnic populations. Additionally, there is limited information available regarding the effectiveness of these obesity assessment methods in Asian countries. In particular, in Asian populations, variations exist in the relative contributions of muscle mass, bone mass and bodily fluids to overall body weight when compared to Europeans. These differences are influenced by cultural subgroups, social and economic conditions, as well as nutritional characteristics [2]. While BFP plays a crucial role in evaluating the prevalence of overweight and obesity, there is a lack of comprehensive data regarding the distribution of body fat in adolescent populations, particularly in Asian regions.

Globally, urbanization significantly propels shifts in dietary patterns and levels of physical activity, with Vietnam standing as no exception to this trend [3]. Eating habits and physical activity levels have been recognized as a key player in changes in BFP [4]. Observations from a study in the city of Tehran showed that a comprehensive lifestyle intervention reduced BFP by 1.81 % in obese adolescents after 12 weeks of intervention [5]. In a study of 764 Indian schoolchildren aged 10–18 years, the authors showed that adolescents with healthy eating habits and regular physical activity had a lower BFP than that in adolescents with negative eating habits and and infrequent physical activity [6]. Besides, a study of 70 schoolchildren aged 14–15 also showed a significant relationship between physical activity and BFP [7]. The analysis of risk factors associated with eating habits and physical activity is the basis for making recommendations and proposing timely measures to reduce the increase in BFP at an early age as well as for proposing measures to prevent overweight and obesity in school age [8–10]. Machine learning is an application of artificial intelligence that helps systems automatically understand data from trained data without needing specific programming. Compared with classical statistics, the new point here is that a machine must efficiently perform inferences and learn from provided data using suitable algorithms and a massive data management facility of a computer. Hence, machine learning is regarded as a pioneering discipline within modern statistics and, more broadly, in the field of data science. Therefore, with the development of big data, use of machine learning algorithms to extract information to produce intelligent data is a necessary task. Research on 7162 Chinese people, using 11 algorithms of machine learning has shown that the random space classification algorithm achieves high overall accuracy and area under the curve. The study, using the evaluation criterion BMI, showed that duration of vigorous intensity activity per week and duration of moderate intensity activity per week were strongly associated with obesity [11]. The study by Babajide et al. (2020) investigated the applicability of machine learning to improve body mass prediction in a dietary intervention program [12]. A study of 40,032 Britons found that machine learning models based on two-dimensional projection allow near-perfect estimation of the volume of adipose tissue [13].

However, at present, studies applying machine learning in analyzing the relationship between eating habits and physical activity to BFP are still very limited. The purpose of this study was to identify risk factors and protective factors for BFP as well as their predictability in the representative population of 11 to 15-year-old students using machine learning techniques. The results of this study are the basis for making recommendations on daily food consumption and physical activity habits as well as developing policies and adjusting intervention programs to control an increase in obesity rates in this population.

Materials and methods. A total of 1208 adolescents (598 boys and 610 girls), aged

11–15 years, originating from nine junior high schools in the capital of Vietnam, were included in the study. Nine schools were selected from 583 junior high schools in Hanoi by simple random sampling method. Students at each school were selected using Epi info 6 software. Healthy students would be eligible for inclusion, if they were free of known chronic disease and were not taking any medication influencing body composition (e.g. β-blockers or diuretics).

The study protocol was approved by the local ethics committee at the Department of Human and Animal Physiology, Faculty of Biology, Hanoi National University of Education. The data collection and storage comply with ethical guidelines and protect individuals' confidentiality. The investigators informed all the school principals, teachers, the participants and their parents for the potential benefits and risks in relation to the study. Parents or guardians provided written informed consent for their child's participation in the study. Throughout the investigation, children had the right to decline answering any questions or to discontinue their participation in the study at any time.

Anthropometric measurements, which included weight, height, waist circumference (WC), and hip circumference (HC), were conducted in accordance with the standardized method developed by National Institute of Nutrition. These measurements were taken with the children wearing minimal clothing, after their shoes and hair ornaments were removed. Height, WC, and HC were measured to the nearest 0.1 cm, while body weight was measured to the nearest 0.1 kg using standardized medical scales. WC was measured at the midpoint between the iliac crest and the lowest rib, and HC was measured at the widest part of the buttocks using an inelastic tape measure. Each of these measurements was performed twice for each child, and the average value was utilized for subsequent analysis. Body Mass Index (BMI) was calculated by dividing the child's weight in kilograms by the square of height in meters. Students' nutritional status was assessed using the criteria established by the International Obesity Task Force (IOTF).

Bioelectrical Impedance Analysis (BIA): participants in this study underwent measurements of body composition using the HBF 375 (Omron) device. Body fat percentage (BFP) was in the focus of this analysis. Prior to these measurements, participants refrained from intense physical activity for at least 12 hours, abstained from consuming food or beverages for at least 3 hours, and ensured they had urinated and defecated within the last 30 minutes. Female participants were excluded from testing during their menstrual period. All BIA measurements were conducted in the morning.

The validation of the survey questionnaire was initially carried out by The National Institute of Nutrition. The questionnaire included questions about food behaviors, main meal frequencies, physical activity, sedentary behaviors, and nutrition knowledge. General information included age, gender, and residence area. Characteristics for eating habits (18 items) included frequencies of eating breakfast; eating speed; number of meals per day; snacking habits; type of food consumed in the snack; time to eat snacks of the day; sensory liking for vegetables / fruits / fatty food / sweet food / carbonated beverage / fast food / animal organs; frequency of consumption of vegetables / fruits / fatty food / sweet food / carbonated beverage / fast food / animal organs. Physical activity was measured based on 27 items: activity preferences, modes of transportation to school, frequency of walking or riding a bicycle, playing or not playing sport games, frequency of exercise, time spent being sedentary, sleep duration, frequency of vigorous, light, or moderate exercise. Nutrition knowledge was collected based on 9 items regarding a general nutrition concept as consuming good food for health or not and obesity-related knowledge as definition of obesity and adequate methods of weight control.

Machine learning algorithms. The decision tree algorithm was used to make predictions for target values in regression tasks or target classes in classification tasks. Our primary emphasis was on solving regression problems, wherein the decision tree's terminal nodes, also known as leaf nodes, can accommodate continuous values, typically in the form of real numbers.

The process of constructing a regression tree involves iteratively dividing the dataset into increasingly smaller subsets while progressively expanding the decision tree structure. The final result is a fully grown tree, comprising decision nodes and leaf nodes. Decision nodes (for example, 'Gender') bifurcate into two or more branches (such as 'Male' and 'Female'), each representing possible attribute values. On the other hand, leaf nodes (for instance, 'BFP value') signify the ultimate decision pertaining to the numerical target value. The topmost decision node in the tree, which corresponds to the most informative predictor, is referred to as the root node. Throughout this experimental investigation, we harnessed the power of the Scikit-learn library, a Python-based machine learning toolkit [14].

Model evaluation. In assessing a Regression Model, we utilize two essential metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) [15]. These metrics serve as indicators of the model's precision and the magnitude of deviation from the actual values. Technically, RMSE is the square root of the mean of the squared errors, whereas MAE is the mean of the absolute errors. In this context, an error represents the disparity between the predicted values (values estimated by our regression model) and the genuine values of a variable. Concretely, the formulas are as follows:

$$RMSE = \sqrt{\frac{\sum (y_i - y_p)^2}{n}}, \qquad (1)$$

$$MAE = \frac{\left|(y_i - y_p)\right|}{n}, \qquad (2)$$

where y_i = actual value; y_p = predicted value, n = number of observations.

Results and discussion. Table 1 shows the characteristics of the study subjects by gender. Boys and girls had similar age, 13.0 and 12.9 years, respectively. However, boys, compared with girls, had a significantly greater weight, height, waist circumference, waist-to-hip

Table 1

General characteristics of the participants

Parameter	Boys (N = 598)	Girls (N = 610)	P-value
Age (years)[b]	13.0 (12.1–13.8)	12.9 (11.9–13.7)	0.406
Weight (kg)[b]	44.8 (37.3–52.2)	43.6 (37.8–49.0)	**0.03**
Height (cm)[a]	154.9 ± 10.3	153.1 ± 6.6	**< 0.0001**
BMI (kg/m^2)[b]	18.5 (16.6–20.8)	18.4 (16.5–20.2)	0.097
Waist circumference (cm)[b]	67.0 (62.5–73.1)	65.3 (62.0–69.6)	**< 0.0001**
Waist to hip ratio[b]	0.81 (0.78–0.86)	0.78 (0.75–0.81)	**< 0.0001**
Nutritional status			
+ Obesity (%)	4.5	0.8	**0.001**
+ Overweight (%)	13.5	11.4	
+ Normal weight (%)	67.6	73.9	
+ Underweight (%)	14.4	13.9	
BFP (%)[b]	18.0 (13.3–22.6)	21.1 (18.4–23.5)	**< 0.0001**
Subcutaneous fat percentage (%)[a]	12.5 ± 4.3	17.9 ± 4.0	**< 0.0001**
Muscle mass percentage (%)[b]	36.3 (33.5–38.7)	29.5 (28.1–31.0)	**< 0.0001**

N o t e: BMI is body mass index; BFP is body fat percentage; [a] data are mean ± SD; [b] data are median (interquartile range). P-values obtained by Student T test or Mann – Whitney U test or Chi-square test. Bold values indicate significant difference between cases and controls.

ratio, and muscle mass percentage ($P < 0.001$). In contrast, the BFP and the subcutaneous fat percentage of boys was lower than that of girls. In boy, BFP was 18.0 %, and subcutaneous fat percentage was 12.5 % meanwhile, in girls, the corresponding values were 21.1 % and 17.9 %, respectively. There was also a difference in nutritional status between male and female adolescents ($P = 0.001$) detected by using the IOTF criteria. Specifically, the rate of overweight and obesity in boys was much higher than in girls (18 % versus 12.2 %).

When investigating the correlations between body fat percentage and some certain anthropometric variables in Vietnamese adolescents, the findings revealed several significant associations. Body fat percentage exhibited strong positive correlations with Body Mass Index ($r = 0.51$), Waist Circumference ($r = 0.42$), and Hip Circumference ($r = 0.31$). Conversely, a negative correlation was observed between body fat percentage and height ($r = -0.30$) (Figure 1).

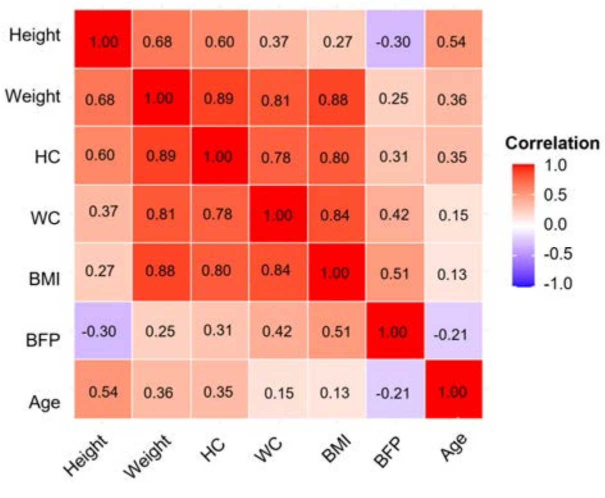

Figure 1. The correlations between body fat percentage and some anthropometric variables

The development of predictive model for adolescents aged between 11 and 15. ***Decision Tree Analysis.*** In assessing a regression model, we utilized two essential metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These metrics serve as indicators of the model's precision and the magnitude of deviation from the actual values. Technically, RMSE is the square root of the mean of the squared errors, whereas MAE is the mean of the absolute errors. In this context, an error represents the disparity between the predicted values (values estimated by our regression model) and the genuine values of a variable.

The experimental results, evaluated using the decision tree model shown in Table 2, indicate RMSE and MAE values of 4.20 and 3.29 for the training dataset, and 4.80 and 3.80 for the validation dataset, respectively. The square root of the average of the squared errors and the average of the absolute errors between the two datasets are relatively small. Therefore, the predictive model developed in this study can be considered highly capable and stable.

Table 2

Decision tree model evaluation ($N = 1208$)

Category	Training (80 %: 966)	Validation (20 %: 242)
RMSE	4.20	4.80
MAE	3.29	3.80

Note: RMSE: Root Mean Squared Error; MAE: Mean Absolute Error.

The decision tree predicted the value of BFP in adolescents from the interaction of environmental factors including gender, pubertal status, and lifestyle-related factors. The decision tree model comprises seven terminal nodes. Notably, among these nodes, four lifestyle-related factors that contribute to an increase in BFP (excluding gender and pubertal status) were as follows: girls who did not consume snacks (predicted BFP was 22.85); girls who both did not consume snacks and had screen time exceeding 2 hours per day (predicted BFP was 21.815); boys who consumed fruit less than three times per week (predicted BFP was 20.749); boys who both consumed fruit less than three times per week and used sedentary mode of transport to schools (predicted BFP was 26.641) (Figure 2).

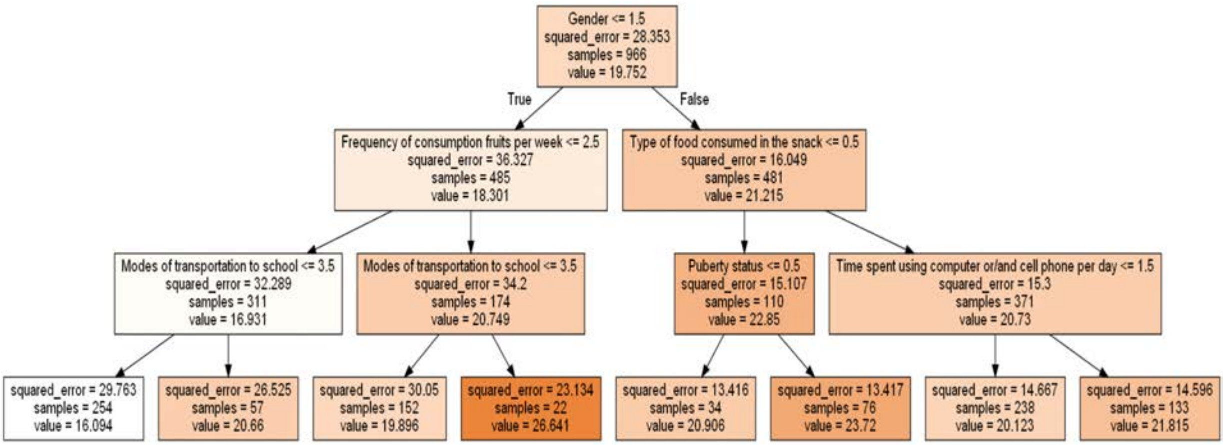

Figure 2. The decision tree model identifies risk groups with high BMI among adolescents aged between 11 and 15: Value is predicted BFP; Gender: 1 = male; 2 = female; Frequency of consumption fruits per week: 1 = > 5 times/week; 2 = 3–5 times/week; 3 = < 3 times/week; Types of food consumed in the snack: 0 = nothing; 1 = milk and dairy products; 2 = high-carb foods; 3 = fast food; Modes of transportation to school: 1 = walking; 2 = cycling; 3 = bus; 4 = motorbike; 5 = car; Puberty status: 0 = Tanner stage 1, 2, 3; 1 = Tanner stage 4, 5; Time spent using computer or / and cell phone per day: 1 = < 2 hours/day; 2 = ≥ 2 hours/day)

Evaluation of the model. Furthermore, to assess the effectiveness of the Decision Tree method more comprehensively, we conducted a comparison with another method, linear regression.

The results on the test data in Figure 1 also reveal that the Decision Tree method exhibits relatively low errors, with RMSE and MAE values of 4.80 and 3.80, respectively. In contrast, linear regression yields higher values of 5.12 for RMSE and 4.20 for MAE (Table 3). Consequently, it can be observed that the Decision Tree method provides more reliable results compared to linear regression.

Table 3

Performance indicators of the machine learning algorithms

Method	RMSE	MAE
Linear regression	5.12	4.20
Decision tree	4.80	3.80

Note: RMSE: Root Mean Squared Error; MAE: Mean Absolute Error.

In this research, a machine learning-centered methodology was employed to predict BFP based on the dietary patterns and physical activity of adolescents aged 11–15 years. The results revealed that within this age group, several lifestyle-related factors exerted the most substantial influence on BFP. These influential factors encompassed the frequency of fruit consumption, snacking behaviors, the mode of transportation to school, and the duration of computer and/or cell phone usage.

Within the context of this research, BIA was employed to measure BFP. BIA finds widespread application within epidemiological and clinical contexts for the assessment of body composition [16]. BIA stands out for its high level of safety, particularly when employed with adolescents and its suitability for widespread use. Despite the inherent margin of error, certain studies indicate that the error of the BIA method remains comparatively minimal [17, 18]. A study conducted on 200 healthy volunteers showed that the correlation coefficient between BIA (using HBF 359) and DXA measurement of BFP was 0.89 ($P < 0.001$) [19].

Our results showed that the Decision Tree method provides more reliable results compared to linear regression. Decision trees represent one of the most widely adopted machine learning algorithms that can effectively address both regression and classification challenges. Decision Tree method stands as a potent instrument for forecasting overweight and obesity, drawing from a diverse array of factors including lifestyle, dietary habits, genetics etc. [20, 21]. Numerous research studies employ a range of machine learning algorithms to attain remarkable accuracy in predicting obesity [22]. Studies using machine learning in the assessment of BFP offer promising insights and have the potential to improve accuracy and efficiency in predicting body fat levels. On the other hand, using decision trees can predict BFP from the aggregate interaction of multiple risk or protective factors instead of just evaluating each factor individually. Accurate BFP assessment through machine learning has significant population health implications [22]. However, prediction of BFP from environmental factors in adolescents is still very limited.

In our study, frequency of fruit consumption and snacking turned out to be the best predictors. Numerous studies have investigated the correlation between diets and BMI, overweight, and obesity among adolescents in various countries. However, the findings from these research endeavors have not been consistently aligned. In a prior study encompassing 34 countries, it was reported that there was no significant association between overweight status and the consumption of fruits or vegetables [23]. Conversely, in a prospective cohort study conducted in the United States involving 14,900 children, it was revealed that intake of fruit and fruit juice did not serve as a predictor for changes in BMI [24]. According to the data from the International Study of Asthma and Allergies in Childhood, encompassing 201,871 adolescents, it was observed that adolescents who consumed fruits, vegetables, pulses, and nuts three or more times per week exhibited a lower BMI in comparison to those who rarely or occasionally consumed these foods [25].

The results of our study also showed that adolescent girls who did not eat snacks had higher BFP than those who did. Numerous research studies indicate that consumption of snacks may lead to an elevation in daily calorie intake instead of diminishing it. Nonetheless, snacking remains an integral component of a wholesome weight loss strategy. Optimal choices for weight loss include snacks that are abundant in complex carbohydrates, protein, and fiber, as they contribute to prolonged satiety [26]. An analysis conducted using data from the China Health and Nutrition Survey, covering the years 2006, 2009, and 2011, indicated that being in the highest snacking tertile was linked to the most substantial reduction in BMI z scores (-2.1) (P < 0.05) in overweight children aged between 7 and 13 years [27].

In this study, it was found that time spent using computer or/and cell phones was important in predicting BFP in adolescents. Some studies have demonstrated that the disproportionate utilization of electronic devices, particularly by adolescents, exerts a significant impact on the physical, psychological, and social well-being of this cohort [28]. An evaluation was conducted on a population-based sample of Finnish twins (N = 4,098); the results indicated that increased time spent using a home computer was linked to a heightened risk of overweight. Moreover, a positive linear trend was observed between cell phone usage and BMI, with a beta coefficient of 0.18 (95 % CI: 0.06–0.30) [29]. The outcomes of the dual-class meta-analysis encompassing 44 studies also revealed that adolescents categorized within the highest range of screen time demonstrated a 1.27-fold increased likelihood of developing overweight or obesity (P < 0.001) [30].

Over the past few decades in Vietnam, the nature of school commuting has shifted towards a more sedentary trend. The existing literature is still inconsistent regarding any associations between school commuting and body composition. The meta-analysis, which encompassed 13 research papers, exclusively concentrated on the connection between active school commuting and the body mass index of children and adolescents. Among the final selection of 13 studies, three studies established definitive associations, three studies indicated partial associations confined to certain subgroups or constrained by societal/geographical factors, while seven studies exhibited no discernible correlations [31].

Strengths and limitations. Some factors associated with BFP in adolescents, as established in prior research, such as physical activity level, were not included in the decision tree model. It's essential to approach these results with caution because the absence of these variables in the model does not necessarily imply that they are unrelated to BFP. Several limitations should be considered in interpreting our findings. Firstly, the generalizability of the results to other ethnic regions may be challenging, as the database used for analysis was specific to adolescents in Hanoi. Secondly, while machine learning can reveal associations, it may not establish causality. Nevertheless, this study boasts several strengths. Firstly, it endeavors to unravel complex relationships among predictor variables using a decision tree analysis approach. Secondly, the application of machine learning techniques allows for more nuanced and data-driven exploration of factors linked to BFP, potentially uncovering non-linear relationships that conventional statistical methods might overlook.

Implications for practice. The findings of this study carry significant practical implications, primarily highlighting the importance of encouraging healthy eating and exercise habits among adolescents. Moreover, parents and schools can incorporate these findings into their childcare strategies. In order to maintain a healthy BFP in children, it is recommended that they consume a minimum of three servings of whole fruits per week and consider incorporating additional healthy snacks alongside their three main meals each day. Additionally, children should prioritize engaging in active modes of transportation while limiting their screen time.

Conclusion. In conclusion, it is noteworthy that this study has successfully identified six distinct groups of predictors related to BFP among adolescents using the decision tree model. Importantly, it took into account multiple variables simultaneously, enhancing the predictive accuracy of the model. Among these predictors, the most significant factors associated with BFP in adolescents include frequency of fruit consumption, snacking habits, mode of transportation to school, and the screen time. This study illuminates some intricate and diverse factors that impact BFP in Vietnamese adolescents. The combination of these factors and interactions with gender and pubertal status can determine BFP in Vietnamese adolescents. However, it's essential to acknowledge that there may still be unclear relationships among these predictors. Therefore, future studies should aim to delve deeper into these associations and further explore the underlying complexities within this multifaceted field of research.

Acknowledgments. The authors would like to express gratitude to Associate Professors Bui Thi Nhung and Tran Quang Binh, as well as colleagues at the National Institute of Nutrition and Hanoi National University of Education, for their generous assistance and support.

Funding. The study was not granted any financial support

Competing interests. The authors declare no competing interests.

References

1. Vanavanan S., Srisawasdi P., Rochanawutanon M., Kumproa N., Kruthkul K., Kroll M.H. Performance of body mass index and percentage of body fat in predicting cardiometabolic risk factors in Thai adults. *Diabetes Metab. Syndr. Obes.*, 2018, vol. 11, pp. 241–253. DOI: 10.2147/DMSO.S167294

2. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet*, 2004, vol. 363, no. 9403, pp. 157–163. DOI: 10.1016/S0140-6736(03)15268-3

3. Mendez M.A., Popkin B.M. Globalization, urbanization and nutritional change in the developing world. *Journal of Agricultural and Development Economics*, 2004, vol. 1, no. 2, pp. 220–241.

4. Bowen L., Taylor A.E., Sullivan R., Ebrahim S., Kinra S., Krishna K.V.R., Kulkarni B., Ben-Shlomo Y. [et al.]. Associations between diet, physical activity and body fat distribution: a cross sectional study in an Indian population. *BMC Public Health*, 2015, vol. 15, pp. 281. DOI: 10.1186/s12889-015-1550-7

5. Kalantari N., Mohammadi N.K., Rafieifar S., Eini-Zinab H., Aminifard A., Malmir H., Ashoori N., Abdi S. [et al.]. Indicator for success of obesity reduction programs in adolescents: body composition or body mass index? Evaluating a school-based health promotion project after 12 weeks of intervention. *Int. J. Prev. Med.*, 2017, vol. 8, no. 73, pp. 128–132. DOI: 10.4103/ijpvm.IJPVM_306_16

6. Madan J., Gosavi N., Vora P., Kalra P. BFP and its correlation with dietary pattern, physical activity and life-style factors in school going children of Mumbai, India. *Obes. Metab. Res.*, 2014, vol. 1, no. 1, pp. 14–19.

7. Dewi R.C., Rimawati N., Purbodjati P. Body mass index, physical activity, and physical fitness of adolescence. *J. Public Health Res.*, 2021, vol. 10, no. 2, pp. 2230. DOI: 10.4081/jphr.2021.2230

8. Steele R.G., Van Allen J. The treatment of pediatric obesity: Bringing contexts and systems into focus. *Children's Health Care*, 2011, vol. 40, no. 3, pp. 171–178. DOI: 10.1080/02739615.2011.590384

9. Briggs D. The roles of managers in addressing Sustainable Development Goals and addressing the burden of chronic disease. *Asia Pacific Journal of Health Management*, 2018, vol. 13, no. 2, pp. 1–3. DOI: 10.24083/apjhm.v13i2.17

10. Alfaleh G., Huffman F.G., Li T., Vaccaro J.A. Child Obesity Prevention Intervention in Kuwaiti Summer Camps Targeting Health Behaviors in Nutrition, Physical Activity, and Screen Time. *Journal of Health Science and Medical Research*, 2021, vol. 39, no. 2, pp. 85–99. DOI: 10.31584/jhsmr.2020765

11. Cheng X., Lin S.-Y., Liu J., Liu S., Zhang J., Nie P., Fuemmeler B.F., Wang Y., Xue H. Does physical activity predict obesity – A machine learning and statistical method-based analysis. *Int. J. Environ. Res. Public Health*, 2021, vol. 18, no. 8, pp. 3966. DOI: 10.3390/ijerph18083966

12. Oladapo B., Tawfik H., Palczewska A., Gorbenko A., Arne A., Martinez J.A., Oppert J.-M., Sørensen T.I.A. A machine learning approach to short-term body weight prediction in a dietary intervention program. *Computational Science – ICCS 2020: 20th International Conference*. Amsterdam, The Netherlands, June 3–5, 2020, pp. 441–455. DOI: 10.1007/978-3-030-50423-6_33

13. Agrawal S., Klarqvist M.D., Diamant N., Stanley T.L., Ellinor P.T., Mehta N.N., Philippakis A., Ng K. [et al.]. Association of machine learning-derived measures of body fat distribution with cardiometabolic diseases in > 40,000 individuals. *MedRxiv*, 2021. DOI: 10.1101/2021.05.07.21256854

14. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Muller A. [et al.]. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011, vol. 12, pp. 2825–2830.

15. Kushwah J.S., Kumar A., Patel S., Soni R., Gawande A., Gupta S. Comparative study of regressor and classifier with decision tree using modern tools. *Materials Today Proceedings*, 2022 vol. 56, no. 6, pp. 3571–3576. DOI: 10.1016/j.matpr.2021.11.635

16. Marra M., Sammarco R., De Lorenzo A., Iellamo F., Siervo M., Pietrobelli A., Donini L.M., Santarpia L. [et al.]. Assessment of body composition in health and disease using bioelectrical impedance analysis (BIA) and dual energy X-ray absorptiometry (DXA): a critical overview. *Contrast Media Mol. Imaging*, 2019, vol. 2019, pp. 3548284. DOI: 10.1155/2019/3548284

17. Cruz Rivera P.N., Goldstein R.L., Polak M., Lazzari A.A., Moy M.L., Wan E.S. Performance of bioelectrical impedance analysis compared to dual X-ray absorptiometry (DXA) in Veterans with COPD. *Sci. Rep.*, 2022, vol. 12, no. 1, pp. 1946–1953. DOI: 10.1038/s41598-022-05887-4

18. Achamrah N., Colange G., Delay J., Rimbert A., Folope V., Petit A., Grigioni S., Déchelotte P., Coëffier M. Comparison of body composition assessment by DXA and BIA according to the body mass index: A retrospective study on 3655 measures. *PLoS One*, 2018, vol. 13, no. 7, pp. e0200465. DOI: 10.1371/journal.pone.0200465

19. Wang J.-G., Zhang Y., Chen H.-E., Li Y., Cheng X.-G., Xu L., Guo Z., Zhao X.-S. [et al.]. Comparison of two bioelectrical impedance analysis devices with dual energy X-ray absorptiometry and magnetic resonance imaging in the estimation of body composition. *J. Strength Cond. Res.*, 2013, vol. 27, no. 1, pp. 236–243. DOI: 10.1519/JSC.0b013e31824f2040

20. Lee I., Bang K.-S., Moon H., Kim J. Risk factors for obesity among children aged 24 to 80 months in Korea: A decision tree analysis. *J. Pediatr. Nurs.*, 2019, vol. 46, pp. e15–e23. DOI: 10.1016/j.pedn.2019.02.004

21. Safaei M., Sundararajan E.A., Driss M., Boulila W., Shapi'i A. A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Comput. Biol. Med.*, 2021, vol. 136, pp. 104754–104780. DOI: 10.1016/j.compbiomed.2021.104754

22. DeGregory K., Kuiper P., DeSilvio T., Pleuss J., Miller R., Roginski J., Fisher C.B., Harness D. [et al.]. A review of machine learning in obesity. *Obes. Rev.*, 2018, vol. 19, no. 5, pp. 668–685. DOI: 10.1111/obr.12667

23. Janssen I., Katzmarzyk P.T., Boyce W.F., Vereecken C., Mulvihill C., Roberts C., Currie C., Pickett W. [et al.]. Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns. *Obes. Rev.*, 2005, vol. 6, no. 2, pp. 123–132. DOI: 10.1111/j.1467-789X.2005.00176.x

24. Field A.E., Gillman M.W., Rosner B., Rockett H.R., Colditz G.A. Association between fruit and vegetable intake and change in body mass index among a large sample of children and adolescents in the United States. *Int. J. Obes. Relat. Disord.*, 2003, vol. 27, no. 7, pp. 821–826. DOI: 10.1038/sj.ijo.0802297

25. Wall C.R., Stewart A.W., Hancox R.J., Murphy R., Braithwaite I., Beasley R., Mitchell E.A., ISAAC Phase Three Study Group. Association between frequency of consumption of fruit, vegetables, nuts and pulses and BMI: analyses of the International Study of Asthma and Allergies in Childhood (ISAAC). *Nutrients*, 2018, vol. 10, no. 3, pp. 316–326. DOI: 10.3390/nu10030316

26. Barnes T.L., French S.A., Harnack L.J., Mitchell N.R., Wolfson J. Snacking behaviors, diet quality, and body mass index in a community sample of working adults. *J. Acad. Nutr. Diet.*, 2015, vol. 115, no. 7, pp. 1117–1123. DOI: 10.1016/j.jand.2015.01.009

27. Taillie L.S., Wang D., Popkin B.M. Snacking is longitudinally associated with declines in body mass index z scores for overweight children, but increases for underweight children. *J. Nutr.*, 2016, vol. 146, no. 6, pp. 1268–1275. DOI: 10.3945/jn.115.226803

28. Kwong C.K.Y., Fong B.Y. Promotion of appropriate use of electronic devices among Hong Kong adolescents. *Asia Pacific Journal of Health Management*, 2019, vol. 14, no. 1, pp. 36–41. DOI: 10.24083/apjhm.v14i1.199

29. Lajunen H.-R., Keski-Rahkonen A., Pulkkinen L., Rose R.J., Rissanen A., Kaprio J. Are computer and cell phone use associated with body mass index and overweight? A population study among twin adolescents. *BMC Public Health*, 2007, vol. 7, pp. 24. DOI: 10.1186/1471-2458-7-24

30. Haghjoo P., Siri G., Soleimani E., Farhangi M.A., Alesaeidi S. Screen time increases overweight and obesity risk among adolescents: a systematic review and dose-response meta-analysis. *BMC Prim. Care*, 2022, vol. 23, no. 1, pp. 161. DOI: 10.1186/s12875-022-01761-4

31. Masoumi H.E. Active Transport to School and Children's Body Weight: A Systematic Review. *TeMA – Journal of Land Use, Mobility and Environment*, 2017, vol. 10, no. 10, pp. 95–110. DOI: 10.6092/1970-9870/4088

Nguyen Thi Hong Hanh, Le Thi Tuyet, Nguyen Thi Trung Thu, Do Thi Nhu Trang, Duong Thi Anh Dao, Le Thi Thuy Dung, Dang Xuan Tho. Identifying the factors related to body fat percentage among Vietnamese adolescents using machine learning techniques. Health Risk Analysis, 2024, no. 1, pp. 158–168. DOI: 10.21668/health.risk/2024.1.16.eng

Received: 25.01.2024
Approved: 29.02.2024
Accepted for publication: 20.03.2024