scientific reports

OPEN

Supervised machine learning algorithms for the classification of obesity levels using anthropometric indices derived from bioelectrical impedance analysis

Rodrigo Yáñez-Sepúlveda[1], Aldo Vásquez-Bonilla[2], Rodrigo Olivares[3], Pablo Olivares[3], Juan Pablo Zavala-Crichton[1], Claudio Hinojosa-Torres[1], Catalina Muñoz-Strale[1], Frano Giakoni-Ramírez[1], Josivaldo de Souza-Lima[1], Jacqueline Páez-Herrera[4], Jorge Olivares-Arancibia[5], Tomás Reyes-Amigo[6], Guillermo Cortés-Roco[7], Juan Hurtado-Almonacid[4], Eduardo Guzmán-Muñoz[8,9], Nicole Aguilera-Martínez[10], José Francisco López-Gil[11,12 ✉], Boryi A. Becerra-Patiño[13], Juan David Paucar-Uribe[13], Exal Garcia-Carrillo[14,15] & Vicente Javier Clemente-Suárez[16,17]

The accurate classification of obesity is essential for public health and clinical decision-making. Traditional anthropometric measures such as body mass index (BMI) have limitations in differentiating between fat and lean mass. This study aimed to evaluate and compare the performance of various supervised machine learning algorithms in classifying obesity levels using anthropometric indices derived from bioelectrical impedance analysis (BIA). A cross-sectional study was conducted on a sample of 5372 adults (age 34.6 ± 10.0 years) (2727 females and 2645 males). Anthropometric data included BMI, fat mass index (FMI), fat-free mass index (FFMI), skeletal muscle index (SMI), muscle mass index (MM), and others were collected using a validated multifrequency octopolar BIA device (InBody 270). Six supervised machine learning models, random forest, gradient koosting, k-nearest neighbors, logistic regression, support vector machine, and decision tree, were trained and evaluated using accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and SHapley Additive exPlanations value explanations. Random forest outperformed all other models, achieving the highest accuracy (84.2%), F1-score (83.7%), and AUC-ROC (0.947). SHapley Additive exPlanations analysis revealed that FMI, FFMI, and BMI were the most influential features, while sex had minimal predictive impact. Machine learning models, particularly tree-based algorithms like random forest, show great potential in classifying obesity levels from anthropometric data with high accuracy and interpretability. These models can enhance the effectiveness of obesity screening in clinical and community settings.

Keywords Machine learning, Bioelectrical impedance analysis, Supervised algorithms, SHAP values, Body composition

[1]Faculty Education and Social Sciences, Universidad Andres Bello, Viña del Mar, Chile. [2]Facultad de Ciencias del Deporte, Universidad de Extremadura, Cáceres, Spain. [3]Escuela de Ingeniería Informática, Universidad de Valparaíso, Valparaíso, Chile. [4]Grupo eFidac, Escuela de Educación Física, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile. [5]Grupo AFySE, Investigación en Actividad Física y Salud Escolar, Escuela de Pedagogìa en Educación Fisica, Facultad de Educación, Universidad de Las Américas, Santiago, Chile. [6]Observatorio de Ciencias de la Actividad Física (OCAF), Departamento de Ciencias de la Actividad Física, Universidad de Playa Ancha, Valparaíso, Chile. [7]Universidad Viña del Mar, Viña del Mar, Chile. [8]Escuela de Kinesiología, Facultad de Salud, Universidad Santo Tomás, Talca, Chile. [9]Escuela de Kinesiología, Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Talca, Chile. [10]Facultad Ciencias de la Salud, Universidad Católica del Maule, Talca, Chile. [11]School of Medicine, Universidad Espíritu Santo, Samborondón, Ecuador. [12]Vicerrectoría de Investigación y Postgrado, Universidad de Los Lagos, Osorno, Chile. [13]Faculty of Physical Education, National Pedagogical University, Bogotá, Colombia. [14]Department of Physical Activity Sciences, Universidad de Los Lagos, Osorno, Chile.

[15]School of Education, Faculty of Human Sciences, Universidad Bernardo O'Higgins, Santiago, Chile. [16]Faculty of Medicine, Health and Sports, Universidad Europea de Madrid, Madrid, Spain. [17]Grupo de Investigación en Cultura, Educación y Sociedad, Universidad de la Costa, Barranquilla, Colombia. ✉email: josefranciscolopezgil@gmail.com

For decades, we have witnessed a steady increase in the prevalence of overweight and obesity worldwide[1]. Obesity is one of the significant public health challenges worldwide, being associated with a substantial increase in the incidence of chronic diseases such as type 2 diabetes, cardiovascular disease and certain cancers[2,3]. In addition, abdominal (central or visceral) obesity is a significant risk factor for cardiovascular disease, diabetes and cancer, and plays a vital role in the metabolic syndrome[4]. Therefore, people living with obesity have different profiles and health needs, but are often referred to as a single entity, defined by a single parameter (i.e., body mass index [BMI]) or not discussed at all[2,5]. Therefore, identifying obesity is crucial for assessing the risk of associated disorders, making it a significant health problem[6]. Simple metrics such as BMI have traditionally been used for diagnosis and classification; however, these do not always accurately reflect actual body composition or metabolic risk[7,8]. BMI is the metric currently used to define anthropometric height/weight characteristics in adults and to classify (categorize) them into groups[9]. Current BMI-based measures of obesity may underestimate or overestimate adiposity and provide inadequate information on individual-level health, undermining medically sound approaches to health care and policy[10]. Furthermore, BMI is insufficient for accurate disease classification of obesity at the individual level because people with similar BMIs often have disparate health risks[11], and do not consider variations in parameters such as body composition, including fat mass (FM), lean mass and lean mass distribution[12,13].

Because of this, there is a need for complementary indices to identify obesity in adults[14]. In this context, derived anthropometric indices, such as (FMI), fat-free mass index (FFMI) and skeletal muscle index (SMI), have demonstrated a greater discriminative ability to characterize different obesity profiles and nutritional status[15–17]. These indices offer a better understanding of obesity by differentiating between FM and fat-free mass (FFM), thus providing a more comprehensive assessment of health status[18]. One of the technologies commonly used to assess body composition, also in clinical trials, is bioelectrical impedance analysis (BIA)[19]. BIA allows the determination of FM and FFM[20,21]. Both FM and lean mass (kg) must be normalized by height squared (m^2), as the FMI and FFMI[22]. Therefore, BIA provides valuable anthropometric data that can help differentiate obesity phenotypes and guide better therapeutic approaches[23].

However, it is important to note that BIA does not constitute a reference method for measuring body composition; methods such as dual X-ray absorptiometry (DXA) are considered reference standards because of their higher accuracy[24]. In addition, BIA is based on predictive models derived from algorithms developed for various populations, which implies that the validity of the results may vary according to the characteristics of the sample evaluated.

In this study, the equipment manufacturer's default algorithm was used, the development and validation of which was performed in a population with similar characteristics but not the same as those of the current sample. Although the instrument has been previously validated against DXA[25], its accuracy has not been confirmed in populations with anthropometric profiles similar to those of the participants in this study, which represents a methodological limitation to be considered. Despite this, given that the aim of the present study was to evaluate the relative predictive value of the BIA-derived indices within a supervised classification model, it was considered that the possible absolute imprecision of the algorithm does not critically compromise the aims of the study, although it should be taken into account when interpreting the results and it is suggested that algorithms be validated for this population to allow the use of more accurate models.

Artificial intelligence (AI) has gained worldwide recognition, including machine learning (ML), which utilizes sophisticated neural networks[26]. AI algorithms can predict obesity; however, more research is needed to evaluate their effectiveness in analyzing obesity-related data and to examine more advanced AI methods[27]. ML can be classified into two main types: unsupervised learning, which operates without labelled data, and supervised learning, which relies on labelled data for guidance[26]. The development of ML algorithms offers an unprecedented opportunity to automate and improve the classification of individuals according to these parameters, facilitating clinical or public health decision-making with greater accuracy and efficiency[27–29]. The use of ML techniques has emerged as a key tool in the detection and management of obesity, allowing the analysis of large volumes of biometric and behavioral data to identify patterns associated with overweight[30,31].

Different studies using ML algorithms to predict obesity have revealed that the use of these models proves highly effective in accurately predicting human obesity[32,33]. By considering various factors, such as demographic information, laboratory results, physical examination findings, and lifestyle factors, these models successfully identify crucial risk factors associated with the obese weight category[34]. The application of ML in this context not only improves the early detection of obesity but also optimizes prevention and treatment strategies, tailoring them to the individual needs of each patient[33,35]. In this sense, ML can transform three significant areas of biomedicine: clinical diagnosis, precision treatments and health monitoring, which aims to maintain health through various diseases and the normal ageing process[36].

This study aimed to compare the performance of various supervised ML algorithms (i.e., random forest, gradient boosting, or support vector machine, among others) in classifying obesity levels in a large sample of adults, based on multiple anthropometric indices obtained through multifrequency bioimpedance. This approach aims to move towards more objective, interpretable and applicable tools in community or clinical settings.

Methods

Participants

A non-experimental, cross-sectional, comparative, and associative study was conducted with a total sample of 5372 adult participants (mean age: 34.6 ± 10.0 years), comprising 2727 women (29.3 ± 8.0 years) and 2645 men (37.8 ± 9.2 years). Participants were recruited through non-probabilistic sampling from a broad spectrum of urban and rural communities, ensuring demographic diversity. Eligibility criteria required individuals to be between 18 and 50 years of age, free from non-communicable chronic diseases, and, where applicable, not pregnant. All subjects provided signed informed consent. Individuals classified as physically active according to the World Health Organization (WHO) guidelines[37], or those with medical conditions affecting body composition, were systematically excluded.

Instrumentation and procedure

All assessments were conducted in standardized environments within community health centers by licensed and certified nursing technicians. Prior to data collection, all technicians underwent rigorous training in the operation of BIA equipment. Participants were thoroughly briefed on the study procedures to ensure full compliance and understanding.

Measurements were obtained under strictly controlled conditions: ambient temperature averaged 20 °C, with a relative humidity of approximately 70%. Participants were instructed to refrain from engaging in intense physical activity, consuming alcohol, or taking diuretics for a minimum of 48 h prior to evaluation. Assessments were conducted following a minimum 4-hour fasting period, with prior gastric and bladder voiding. Subjects were measured barefoot and in undergarments, with all contact areas disinfected using 70% isopropyl alcohol in accordance with sanitation protocols.

A validated multi-frequency, octopolar bioelectrical impedance analyzer (InBody 270) was employed to assess body composition. Parameters extracted via the Lookin' Body software suite included: body weight, height, BMI, fat mass (kg and %), FFM, total body water (TBW), skeletal muscle mass (SMM), and basal metabolic rate (BMR), among others.

Anthropometric indices

To comprehensively evaluate and classify adiposity and body composition, the following anthropometric indices were calculated:

- BMI: Weight (kg) divided by height squared (m^2).
- FMI: FM (kg) divided by height squared (m^2).
- FFMI: FFM adjusted for height.
- SMI: Appendicular muscle mass (kg) divided by height squared (m^2).
- Muscle mass index (MMI): Total muscle mass (kg) divided by squared height (m^2).

Nutritional status was classified according to the thresholds proposed by Harty et al.[38], adapted to general population parameters.

Statistical analysis

All statistical analyses were conducted using jamovi software (version 2.3.21). Descriptive data were expressed as mean ± standard deviation. Between-group comparisons by sex were performed using contingency tables and inferential statistics. The assumption of normality was assessed via the Shapiro-Wilk test, followed by analysis of variance (ANOVA) and Bonferroni post hoc corrections for pairwise comparisons. Effect size was estimated using Cohen's d, with interpretation thresholds defined as follows: small (≥ 0.2–0.5), medium (> 0.5–0.8), and large (> 0.8)[39]. Statistical significance was set at $p < 0.05$.

Machine learning analysis

The ML analysis was performed in Jupyter Notebook, and the Python programming language was used to develop codes.

1. Data acquisition

A multidimensional dataset was compiled, integrating anthropometric parameters and lifestyle-related variables derived from structured questionnaires and physical measurements. Variables included weight, height, BMI, age, sex, physical activity level, caloric intake, and other health-related metrics. Rigorous data cleaning protocols were applied to remove incomplete entries and outliers, ensuring dataset integrity. Before training the models, outliers were removed to optimize the quality of the analysis using z-scores.

2. Data preprocessing

Prior to model training, a series of preprocessing steps were implemented:

Numerical normalization using minimum-maximum (min-max) scaling.
One-hot encoding for categorical variables such as gender and dietary habits.
The dataset was split into training (70%) and testing (30%) sets using stratified sampling, preserving the proportional distribution of obesity categories.

3. Feature selection

To enhance model efficiency and interpretability, both statistical and algorithmic feature selection techniques were employed. Correlation analysis and recursive feature elimination with a random forest base estimator were utilized to identify and retain the most predictive variables while minimizing dimensionality.

4. Supervised classification algorithms

Multiple supervised ML algorithms were trained and benchmarked to classify individuals into distinct obesity categories. The models included:

– Support vector machine with radial basis function kernel.
– Random forest classifier with hyperparameter tuning.
– K-nearest neighbors with optimal k selection.
– Logistic regression.
– Gradient Boosting for high-performance ensemble modeling.
– Decision tree using recursive binary partitioning based on impurity measures (e.g., Gini index or information gain), producing a hierarchical, tree-structured model for classification or regression tasks.

5. Model performance evaluation

Models were assessed using a comprehensive suite of classification metrics:

– Accuracy, precision, recall, and F1-score.
– Confusion matrix to examine class-level performance.
– 5-fold cross-validation to ensure generalizability.
– Area under the receiver operating characteristic curve (AUC-ROC).
– Feature importance plots (for tree-based models) to interpret variable contributions from SHapley Additive exPlanations (SHAP).

6. Final implementation and internal validation on a held-out test set.

The best-performing model was validated on an independent test set, evaluating its robustness and generalization capability. The final model's implications were discussed in terms of its clinical applicability and potential for integration into decision support systems in public health and preventive medicine frameworks.

Ethical considerations

This study was conducted in full compliance with the ethical principles outlined in the Declaration of Helsinki[40] and the International Ethical Guidelines for Health-related Research Involving Humans issued by the Council for International Organizations of Medical Sciences (CIOMS)[41]. The evaluation protocols were approved by the Scientific Ethics Committee of the Universidad Viña del Mar (Code R62- 19a). Prior to data collection, the study protocol received approval from the appropriate institutional ethics review board. All participants were thoroughly informed about the study's aims, procedures, potential risks, and benefits, and provided written informed consent before inclusion. Confidentiality and anonymity were maintained by assigning coded identifiers and securing data in password-protected files accessible only to authorized personnel. Participants were informed of their right to withdraw from the study at any time without consequences. Furthermore, all assessments were conducted under standardized and safe conditions, with trained healthcare professionals present to ensure participants' well-being and adherence to ethical protocols.

Results

Table 1 presents descriptive statistics (mean ± standard deviation) for body composition indices stratified by fat classification (normal, high, very high). Individuals in the "very high" fat category showed the highest BMI mean (29.7 ± 3.0) and FMI (9.8 ± 2.4), whereas the "normal" fat group exhibited lower values in both parameters (BMI: 25.4 ± 2.5; FMI: 5.6 ± 1.6). Notably, the SMI and muscle mass index remained relatively stable across groups, with minimal variation. Similarly, the FFMI showed marginal differences between categories. These findings suggest that increases in fat classification are primarily associated with adiposity-related measures, while lean mass components remain comparatively consistent.

Six supervised learning algorithms were compared for a multiclass classification (three classes: normal, high, and very high), with their performance evaluated in terms of accuracy, precision, recall, F1-score, training time, and AUC-ROC. The results are summarized in Table 2.

Fat Classification	Body mass index	Skeletal muscle index	Fat mass index	Muscle mass index	Fat-free mass index
High	28.1 ± 3.3	8.3 ± 0.9	9.1 ± 2.3	11.2 ± 1.3	19.3 ± 0.9
Normal	25.4 ± 2.5	8.5 ± 1.0	5.6 ± 1.6	11.4 ± 1.4	19.8 ± 1.1
Very high	29.7 ± 3.0	8.4 ± 0.9	9.8 ± 2.4	11.3 ± 1.3	19.6 ± 0.9

Table 1. Comparison of the various indexes used according to the level of obesity classification. Results are shown in mean and standard deviation.

Model	Accuracy	Precision	Recall	F1 score	Train time (s)	AUC-ROC
Random forest	0.842	0.836	0.842	0.837	0.718	0.947
Gradient boosting	0.815	0.805	0.815	0.801	1.988	0.932
K-nearest neighbors	0.813	0.806	0.813	0.808	0.003	0.910
Support vector machine	0.696	0.543	0.696	0.609	1.508	0.879
Logistic regression	0.751	0.714	0.751	0.707	0.058	0.827
Decision tree	0.813	0.815	0.813	0.814	0.017	0.832

Table 2. Comparison of the quality of the machine learning models. *AUC-ROC*, area under the receiver operating characteristic curve.
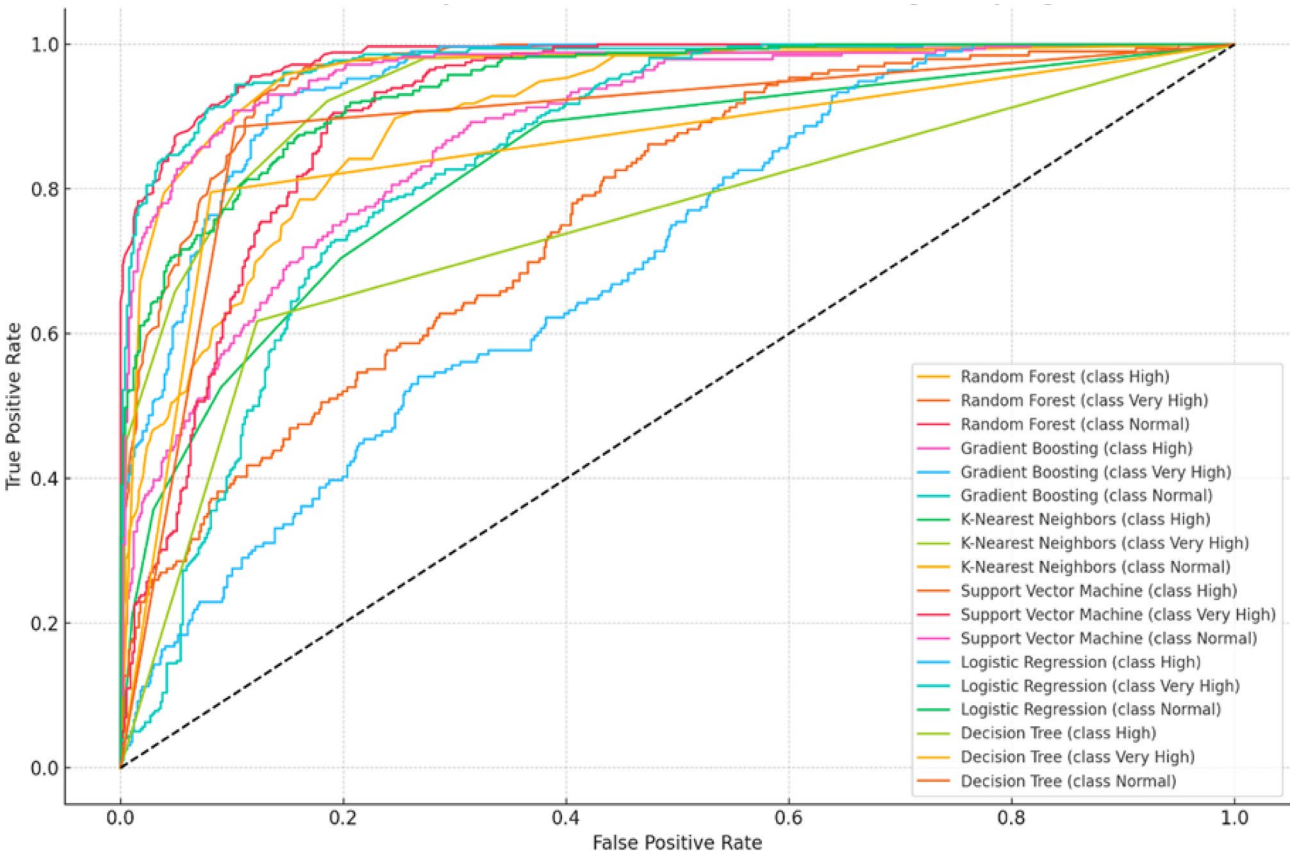
Fig. 1. Comparison of area under the receiver operating characteristic curves across machine learning models and obesity categories (normal, high, very high).

The results presented in Table 2 indicate that the random forest model achieved the best overall performance, with an accuracy of 84.21%, a precision of 83.66%, a recall of 84.21%, and an F1-score of 83.74%, outperforming all other evaluated models. Gradient boosting and k-nearest neighbors also demonstrated competitive performance, each yielding F1-scores above 80%. In contrast, the support vector machine model exhibited the lowest performance, with an F1-score of only 60.96%, highlighting its limited ability to handle the complexity of the classification task. These findings support the selection of tree-based models, particularly random forest, as the most suitable approach for the multiclass classification task in this study, due to their robustness, stability, and ability to generalize effectively in the presence of nonlinear relationships.

In Fig. 1 the AUC-ROC curves by model and class reveal that tree-based models, particularly random forest and gradient boosting, exhibit superior discriminative capacity across the three evaluated classes, consistently approaching the upper-left corner of the graph. This pattern reflects a high true positive rate coupled with a low false positive rate, indicative of excellent classification performance. In contrast, the support vector machine model displays noticeably flatter curves, especially for classes 0 and 1, suggesting a limited ability to correctly distinguish between classes. K-nearest neighbors, logistic regression, and decision tree show intermediate performance, with moderately high AUC-ROC curves for some classes but lacking the consistency and robustness observed in ensemble methods. Collectively, these results reinforce the conclusion that random
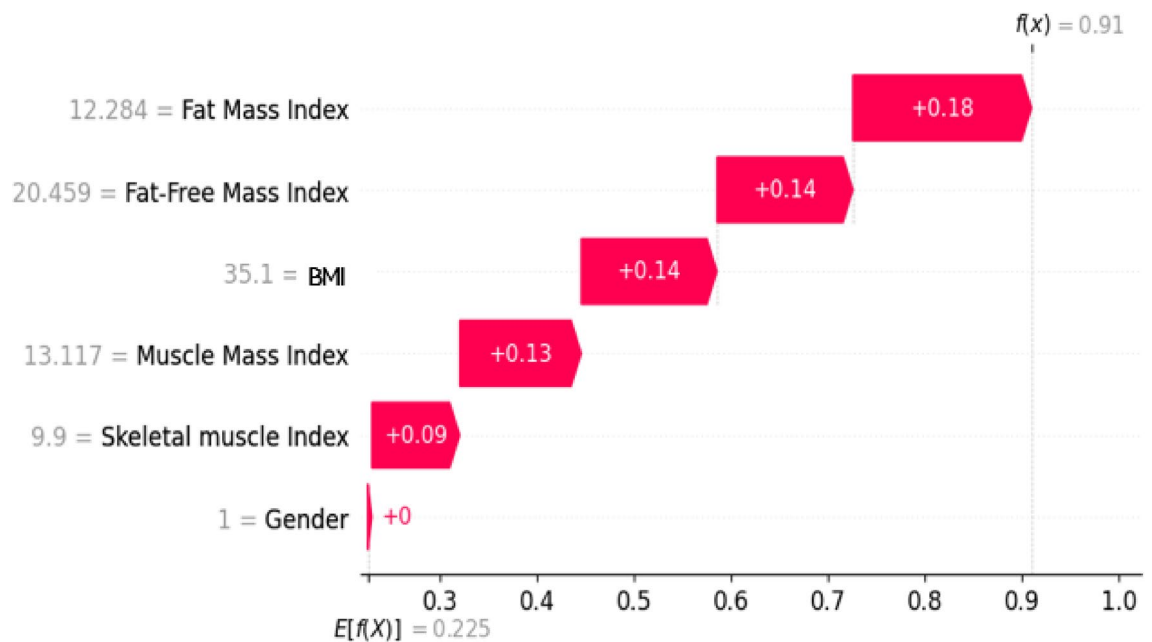
Fig. 2. SHapley Additive exPlanations value summary of feature contributions to obesity classification in a single prediction. *BMI*, body mass index.
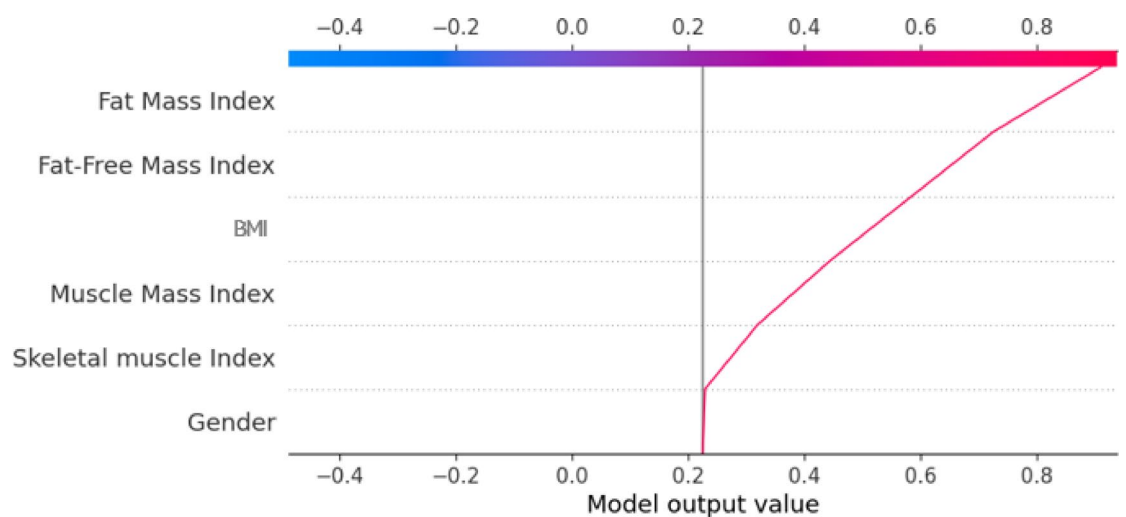

Fig. 3. Cumulative SHapley Additive exPlanations value plot showing feature impact on model output for obesity prediction.

forest provides the best balance between sensitivity and specificity in a multiclass context, further supporting its position as the most robust and generalized model for classifying fat level categories.

Figure 2 presents a local SHAP explanation for a specific model prediction, illustrating how the final output $f(x) = 0.91$, starting from a baseline value of $E[f(x)] = 0.225$, results from the cumulative contributions of various input features. The FMI emerges as the primary positive driver, contributing + 0.18 to the final probability, followed by the FFMI, BMI, and MMI, each contributing between + 0.13 and + 0.14. The SMI also exerts a positive influence (+ 0.09), while gender shows no significant impact. This analysis highlights that the high predicted probability is predominantly driven by features related to body composition, particularly fat and muscle mass, thereby supporting the physiological validity of the model's behavior in classifying fat level categories.

Figure 3 shows a cumulative SHAP graph illustrating how each feature progressively contributes to the model's prediction of a specific instance. The final model output converges to approximately 0.91, starting from a baseline value of around 0.225, which represents the model's average output in the absence of individualized input. The strongest contributor to the elevated prediction is the FMI, followed by the FFMI, BMI, muscle mass index, and SMI, all of which incrementally drive the prediction upward. Gender appears once again as a neutral
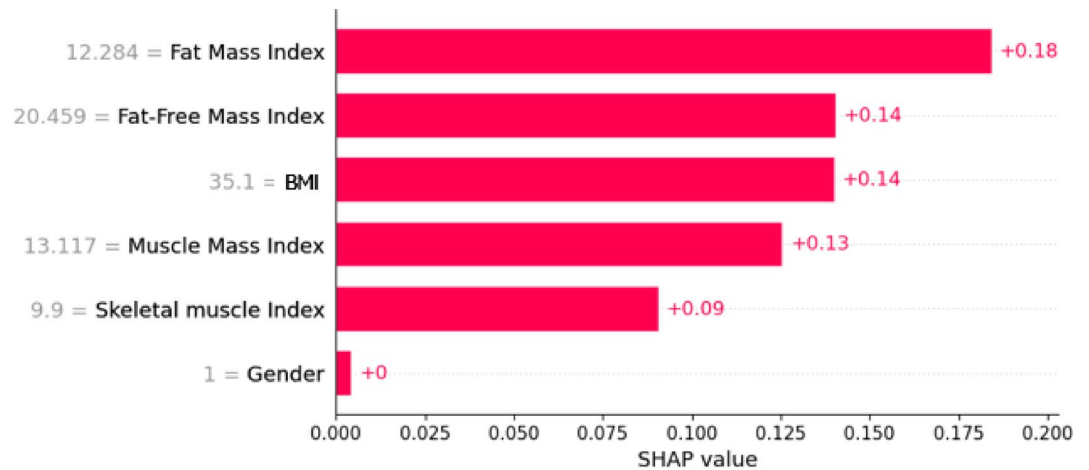
Fig. 4. Feature importance ranked by SHapley Additive exPlanations values for a single prediction in obesity modeling. *BMI*, body mass index; *SHAP*, SHapley Additive exPlanations.

Fig. 5. SHapley Additive exPlanations summary plot depicting feature impact and value distribution on obesity classification. *BMI*, body mass index; *SHAP*, SHapley Additive exPlanations.

variable, with no meaningful impact. A color transition from blue (indicating negative impact and low values) to red (positive impact, high values) visually confirms that higher values in these physiological features are strongly associated with the predicted classification. This plot reinforces the interpretation that body composition is the primary determinant in the model's output and that the model aligns with biologically grounded metrics in predicting body fat levels.

The SHAP bar chart in Fig. 4 shows the individual contributions of the features to a specific prediction, ranked by importance. The FMI stands out as the most influential variable, contributing + 0.18 to the final classification probability, followed by the FFMI and BMI, each contributing + 0.14. These values suggest that the model relies not only on the absolute fat content but also on the balance between fat and lean mass in making its decision. Next in importance are the MMI (+ 0.13) and the SMI (+ 0.09), indicating that muscularity also plays a role in the model's output. Gender shows no contribution, suggesting that the model does not exhibit bias toward this variable in this instance. Overall, this visualization reinforces the physiological coherence of the model: it predicts higher body fat levels when confronted with elevated values in key indicators of total body mass, fat, and muscle.

This SHAP bee diagram illustrated in Fig. 5 shows the impact of each feature on the model´s output across multiple instances, capturing both the magnitude and direction of influence. The FMI emerges as the most critical variable, with high values (shown in red) associated with positive SHAP values, thus increasing the model's predicted probability. In contrast, low values (blue) correspond to negative contributions. A similar, though less pronounced, pattern is observed for the FFMI, BMI, and muscle mass index, all of which show moderate dispersion, suggesting a consistent and physiologically plausible influence on the model's predictions. In contrast, gender and SMI display low impact, with SHAP values clustered around zero, indicating minimal or negligible overall effect on the classification outcome.
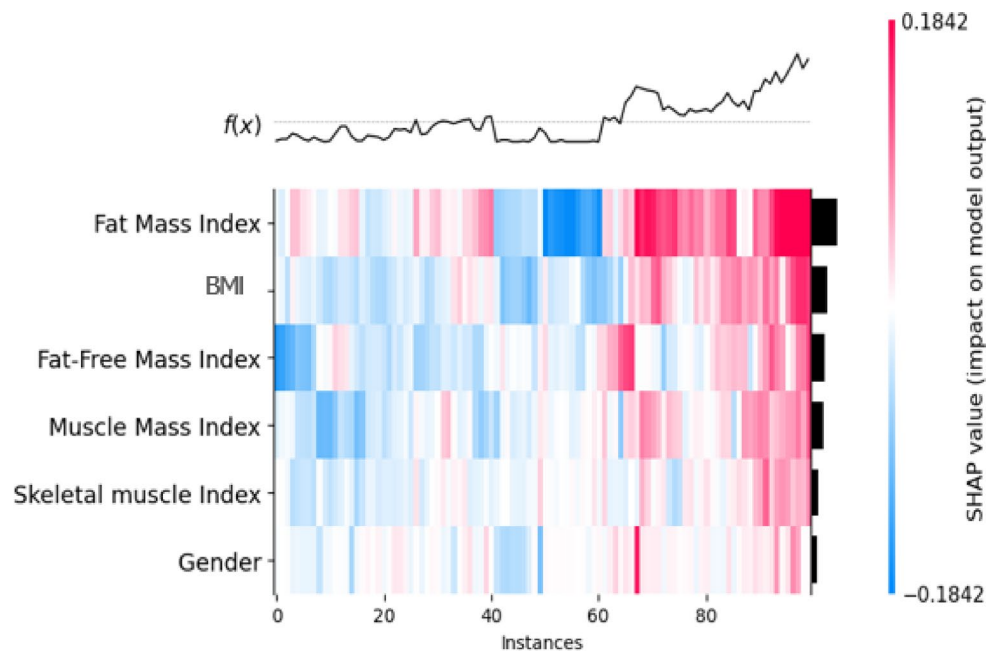
Fig. 6. SHapley Additive exPlanations heatmap showing feature contributions across individuals ordered by model output in obesity prediction. BMI, body mass index; *SHAP*, SHapley Additive exPlanations.

The SHAP heatmap in Fig. 6 visualizes the cumulative impact of each feature on model predictions across multiple instances. The top panel displays the variation in the model's output f(x), while the heatmap below encodes the SHAP values of individual features using a colour scale, blue indicating negative contributions and red indicating positive ones. The FMI stands out as the most influential variable, with consistently strong positive effects (intense red) in instances where the model output is high. It is followed in importance by BMI, FFMI, and MMI, all of which exhibit increasing contributions in higher predictions, albeit with smaller magnitude. In contrast, Gender and SMI display near-neutral effects, with SHAP values close to zero. This visualization reinforces the conclusion that the model's predictions are primarily driven by body composition variables, especially FM, and that its behavior remains consistent and physiologically plausible across the evaluated population.

Discussion

The results of the present study show that tree-based decision models, notably the random forest algorithm, performed best in classifying obesity levels, with superior values for accuracy (83.6%) and AUC-ROC (0.947). Another study[42] evaluated the effectiveness of non-dietary factors, such as lifestyle, family history and demographic characteristics, in predicting obesity using ML models. Analyzing data from more than 2,000 individuals aged 14–61 years, several algorithms were tested, with random forest being the most accurate, with an AUC-ROC of 92.3% and an accuracy of 66.9%, demonstrating that it is possible to detect obesity without resorting to dietary data, which may facilitate more accessible and earlier preventive assessments[43]. Similar results were found by, who developed a prediction model for obesity levels using nine ML algorithms[43]. The results showed that the random forest algorithm performed the best, with an accuracy of 92.29%. Also, Dirik[44] showed that random forest achieved the highest accuracy with 95.78%, while logistic regression followed closely with 95.22%. These findings are consistent with previous research indicating the robustness of ensemble models to non-linear relationships and multivariate data in public health[45,46].

Moreover, the relative importance of the predictor variables, assessed by SHAP values, indicated that the most influential indices were FMI, FFMI and BMI. This result reaffirms the relevance of directly measuring body composition rather than relying exclusively on BMI to characterize excess fat[47,48]. Studies have shown that FMI provides a more accurate measure of adiposity and is more strongly associated with health risks such as hypertension, type 2 diabetes and cardiovascular disease compared to BMI alone[49–51]. In a previous study, Górnika et al.[16], according to their results, point out that FMI and FM percentage could be considered the best markers for the detection of obesity in adults, independent of sex. The low impact of sex as a predictor variable in our study is also in line with studies reporting that height-adjusted body composition can eliminate gender-related biases[52]. Furthermore, cross-sectional research has found that a high FMI is positively associated with a higher prevalence of metabolic syndrome independently of BMI and body fat percentage[53].

The SHAP value is a uniform measure of the importance of features used in ML models[54,55]. Therefore, the use of explanatory tools such as SHAP adds significant value to ML models in healthcare by providing insight into how and why a given classification occurs, improving clinical confidence and algorithmic transparency[56]. Furthermore, it solves the problem of poor readability, better interpreting the model established by ML and applying it to early detection, monitoring and intervention of obesity[57]. In the study by Lin et al.[57], although

the SHAP value was used to visualize the effects of features in the model, their results showed that waist circumference had a positive impact on the predictive power and female sex was a positive predictor of obesity, but the sample consisted of individuals with overweight.

It is striking that in our study, the SMI showed almost neutral effects as a predictor of obesity, with SHAP values close to zero, even below BMI and FFM. These results are unexpected considering that skeletal muscle plays a key role in health and metabolic efficiency[58]. About FFM in our study it had a SHAP value of + 14 following in importance of contribution to FMI as a predictor of obesity, considering that higher FFMI is related to better physical fitness, higher metabolic rate and better overall health outcomes, suggesting that individuals with a higher ratio of FFM (lean tissue such as muscle and bone) to their FM may indicate a more favourable fat distribution[59], which is associated with better metabolic health, as lean mass, particularly muscle mass, is associated with a lower risk of metabolic syndrome, better insulin sensitivity and better overall physical function[60]. This highlights the importance of considering both fat and lean mass when assessing body composition, as different proportions of fat and lean tissue may have other implications for health[59,61,62].

This study makes a significant contribution to the field of health and AI applied to diagnosis by demonstrating the effectiveness of supervised ML algorithms for classifying obesity levels based on anthropometric indices obtained through BIA. The main contributions include:

- Validation of the BIA-based approach: it was demonstrated that bioelectrical impedance-derived indices, such as the FMI, FFMI, and BMI, are highly predictive variables for classifying obesity levels. This supports the use of BIA as a non-invasive and efficient tool in clinical and community settings.
- Identification of the best predictive model: When comparing different supervised algorithms, the random forest model stood out as the most effective, achieving an accuracy of 84.2%, an F1 score of 83.7% and an AUC-ROC value of 0.947. This demonstrates its robustness and reliability for this type of classification task.
- Application of model interpretability: By analyzing SHAP values, the study identified the most influential variables in the model's predictions, highlighting the greater predictive value of body composition indices compared to demographic variables such as sex, which showed minimal predictive impact. This transparent interpretation facilitates confidence in the model among health professionals.
- Contribution to personalized and preventive medicine: By establishing an accurate approach to classifying obesity levels using individual variables derived from BIA, the study offers a potential tool to support personalized clinical decisions aimed at the prevention and early management of overweight and obesity.
- Advancing the integration of AI in healthcare: The work represents a breakthrough in the effective integration of AI techniques in the biomedical field, demonstrating that supervised models can not only automate classification tasks but also improve understanding of the factors underlying complex conditions such as obesity.

Since BIA-derived anthropometric indices are obtained rapidly and with minimal operational burden, the values could be integrated in an automated fashion into electronic medical records to generate real-time risk stratification, trigger clinical decision alerts (e.g., referral for nutritional counseling, metabolic assessment, or intensive follow-up), and support tiered resource allocation. At the public health level, the algorithm could be incorporated into screening campaigns in schools, community primary care or workplace settings, provided that standardized measurement protocols, calibration between devices and action thresholds adapted by age, sex and epidemiological context are established. It is essential to consider data interoperability, multicenter external validation prior to deployment, cost-effectiveness analysis and equity surveillance (to ensure that historically underserved populations are not excluded or misclassified). An expanded discussion along these lines would situate the findings not only as an algorithmic proof of concept, but as a potential practical component of early obesity prevention and management networks.

Limitations

There are some limitations to this study. First, although a large sample was used, the sampling method was not probabilistic, which limits the generalizability of the results to the entire population. Second, anthropometric data were collected at a single point in time, preventing the assessment of longitudinal changes. In addition, although the algorithms were evaluated by cross-validation, external studies in other populations and contexts would be required to verify their general applicability. The model did not include metabolic or clinical variables (e.g. lipids or glucose), which could have enriched the prediction, considering that obesity is a heterogeneous clinical entity with distinct subtypes based on genetic architecture and phenotypic biomarkers including measures of insulin sensitivity, glycaemia, fitness, body composition and cardiovascular risk[63]. Finally, although the study's limitations are acknowledged, a more in-depth reflection on potential bias arising from the use of non-probabilistic sampling would be warranted, as this approach may have led to the overrepresentation or underrepresentation of certain population subgroups. Finally, despite the relatively large sample size, the results were not stratified considering (e.g., age group, sex, socioeconomic status, or race/ethnicity), further limiting the generalizability of the findings to other populations or settings.

Conclusions

This study demonstrates that supervised ML algorithms, particularly random forest, are effective and accurate tools for classifying obesity levels from multiple anthropometric indices. The incorporation of explanatory models such as SHAP allows for a clear interpretation of the factors influencing the classification, promoting a safer and more understandable application in clinical or public health contexts. The multivariate and interpretable model-based approach represents a relevant step towards personalized medicine, where decisions on nutritional diagnosis can be based on more complex and representative models than traditional BMI.

Data availability

Data will be made available on request, link [https://figshare.com/articles/dataset/DATA_BIA_SHAP/2961374 0?file=56433842].

Code availability

https://github.com/rigo1983/Code-BIA/blob/main/Codigo%20SHAP%20SR.

References

1. Sørensen, T. Forecasting the global obesity epidemic through 2050. *Lancet* **405**(10481), 756–757. https://doi.org/10.1016/S0140-6736(25)00260-0 (2025).
2. Bluher, M. Obesity: Global epidemiology and pathogenesis. *Nat. Rev. Endocrinol.* **15**(5), 288–298. https://doi.org/10.1038/s41574-019-0176-8 (2019).
3. Piqueras, P. et al. Anthropometric indicators as a tool for diagnosis of obesity and other health risk factors: A literature review. *Front. Psychol.* **12**, 631179. https://doi.org/10.3389/fpsyg.2021.631179 (2021).
4. The Lancet Diabetes Endocrinology. Redefining obesity: Advancing care for better lives. *Lancet Diabetes Endocrinol.* **13**(2), 75. https://doi.org/10.1016/S2213-8587(25)00004-X (2025).
5. Zhou, X. et al. Association of anthropometric and obesity indices with abnormal blood lipid levels in young and middle-aged adults. *Heliyon* **11**(1), e41310. https://doi.org/10.1016/j.heliyon.2024.e41310 (2024).
6. Nimptsch, K., Konigorski, S. & Pischon, T. Diagnosis of obesity and use of obesity biomarkers in science and clinical medicine. *Metabolism* **92**, 61–70. https://doi.org/10.1016/j.metabol.2018.12.006 (2021).
7. Frühbeck, G. et al. Obesity: The gateway to ill health—an EASO position statement on a rising public health, clinical and scientific challenge in Europe. *Obes. Facts* **6**(2), 117–120. https://doi.org/10.1159/000350627 (2013).
8. Rubino, F. et al. Definition and diagnostic criteria of clinical obesity. *Lancet Diabetes Endocrinol.* **13**(3), 221–262. https://doi.org/10.1016/S2213-8587(24)00316-4 (2025).
9. Coral, D. E. et al. Subclassification of obesity for precision prediction of cardiometabolic diseases. *Nat. Med.* **31**(2), 534–543. https://doi.org/10.1038/s41591-024-03299-7 (2025).
10. Ceniccola, G. D. et al. Current technologies in body composition assessment: Advantages and disadvantages. *Nutrition* **62**, 25–31. https://doi.org/10.1016/j.nut.2018.11.028 (2019).
11. Carbone, S., Lavie, C. J. & Arena, R. Obesity and heart failure: Focus on the obesity paradox. *Mayo Clin. Proc.* **92**(2), 266–279. https://doi.org/10.1016/j.mayocp.2016.11.001 (2017).
12. Merchant, R. A. et al. Relationship of fat mass index and fat free mass index with body mass index and association with function, cognition and sarcopenia in Pre-Frail older adults. *Front. Endocrinol.* **12**, 765415. https://doi.org/10.3389/fendo.2021.765415 (2021).
13. Kim, C. H. et al. Norm references of fat-free mass index and fat mass index and subtypes of obesity based on the combined FFMI-%BF indices in the Korean adults aged 18–89 year. *Obes. Res. Clin. Pract.* **5**(3), e169–e266. https://doi.org/10.1016/j.orcp.2011.01.004 (2011).
14. Romero-Corral, A. et al. Accuracy of body mass index in diagnosing obesity in the adult general population. *Int. J. Obes.* **32**(6), 959–966. https://doi.org/10.1038/ijo.2008.11 (2008).
15. Gažarová, M., Bihari, M., Lorková, M., Lenártová, P. & Habánová, M. The use of different anthropometric indices to assess the body composition of young women in relation to the incidence of obesity, sarcopenia and the premature mortality risk. *Int. J. Environ. Res. Public Health* **19**(19), 12449. https://doi.org/10.3390/ijerph191912449 (2022).
16. Górnicka, M. et al. Anthropometric indices as predictive screening tools for obesity in adults; the need to define Sex-Specific Cut-Off points for anthropometric indices. *Appl. Sci.* **12**(12), 6165. https://doi.org/10.3390/app12126165 (2022).
17. Gažarová, M., Bihari, M. & Šoltís, J. Fat and fat-free mass as important determinants of body composition assessment in relation to sarcopenic obesity. *Rocz. Panstw. Zakl. Hig.* **74**(1), 59–69. https://doi.org/10.32394/rpzh.2023.0243 (2023).
18. Khalil, S. F., Mohktar, M. S. & Ibrahim, F. The theory and fundamentals of bioimpedance analysis in clinical status monitoring and diagnosis of diseases. *Sensors* **14**(6), 10895–10928. https://doi.org/10.3390/s140610895 (2014).
19. Bosy-Westphal, A. & Müller, M. J. Diagnosis of obesity based on body composition-associated health risks—Time for a paradigm change. *Obes. Rev.* **22**(2), e13190. https://doi.org/10.1111/obr.13190 (2021).
20. Salihefendic, N., Zildzic, M., Masic, I. & Jankovic, S. M. Anthropometric data by using bioelectrical analysis as parameters for new classification and definition of obesity. *Mater. Soc. Med.* **37**(1), 11–17. https://doi.org/10.5455/msm.2024.37.11-17 (2025).
21. Genc, A. C. & Arıcan, E. Obesity classification: A comparative study of machine learning models excluding weight and height data. *Rev. Assoc. Med. Bras.* **71**(1), e20241282 (2025).
22. Rostam Niakan Kalhori, S., Najafi, F., Hasannejadasl, H. & Heydari, S. Artificial intelligence-enabled obesity prediction: A systematic review of cohort data analysis. *Int. J. Med. Inf.* **196**, 105804. https://doi.org/10.1016/j.ijmedinf.2025.105804 (2025).
23. Kehinde, O. Machine learning in predictive modelling: Addressing chronic disease management through optimized healthcare processes. *Int. J. Res. Publ. Rev.* **6**, 1525–1539 (2025).
24. Scafoglieri, A. & Clarys, J. P. Dual energy X-ray absorptiometry: Gold standard for muscle mass? *J. Cachexia Sarcopenia Muscle* **9**(4), 786–787. https://doi.org/10.1002/jcsm.12308 (2018).
25. Ballesteros-Pomar, M. D. et al. Bioelectrical impedance analysis as an alternative to dual-energy x-ray absorptiometry in the assessment of fat mass and appendicular lean mass in patients with obesity. *Nutrition* **93**, 111442. https://doi.org/10.1016/j.nut.2021.111442 (2022).
26. Esteva, A. et al. A guide to deep learning in healthcare. *Nat. Med.* **25**(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z (2019).
27. Azmi, S. et al. Harnessing artificial intelligence in obesity research and management: A comprehensive review. *Diagnostics* **15**(3), 396. https://doi.org/10.3390/diagnostics15030396 (2025).
28. Huang, L. et al. The role of artificial intelligence in obesity risk prediction and management: Approaches, insights, and recommendations. *Medicina* **61**(2), 358. https://doi.org/10.3390/medicina61020358 (2025).
29. Choong, C. et al. Identifying individuals at risk for weight gain using machine learning in electronic medical records from the united States. *Diabetes Obes. Metab.* **27**(6), 3061–3071. https://doi.org/10.1111/dom.16311 (2025).
30. Jawara, D. et al. Using machine learning to predict weight gain in adults: an observational analysis from the all of Us research program. *J. Surg. Res.* **306**, 43–53. https://doi.org/10.1016/j.jss.2024.11.042 (2025).
31. Huang, A. A. & Huang, S. Y. Application of a transparent artificial intelligence algorithm for US adults in the obese category of weight. *PLoS One* **19**(5), e0304509. https://doi.org/10.1371/journal.pone.0304509 (2024).
32. Atkinson, J. G. & Atkinson, E. G. Machine learning and health care: Potential benefits and issues. *J. Ambul. Care Manag.* **46**(2), 114–120. https://doi.org/10.1097/JAC.0000000000000453 (2023).

33. Bays, H. E. et al. Artificial intelligence and obesity management: an obesity medicine association (OMA) clinical practice statement (CPS) 2023. *Obes. Pill.* **6**, 100065. https://doi.org/10.1016/j.obpill.2023.100065 (2023).

34. 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.* **136**, 104754 (2021).

35. An, R., Shen, J. & Xiao, Y. Applications of artificial intelligence to obesity research: Scoping review of methodologies. *J. Med. Internet Res.* **24**(12), e40589. https://doi.org/10.2196/40589 (2022).

36. Goecks, J., Jalili, V., Heiser, L. & Gray, J. W. How machine learning will transform biomedicine. *Cell* **181**(1), 92–101. https://doi.org/10.1016/j.cell.2020.03.022 (2020).

37. Bull, F. C. et al. World health organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* **54**(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955 (2020).

38. Harty, P. S. et al. Military body composition standards and physical performance: Historical perspectives and future directions. *J. Strength. Cond Res.* **36**(12), 3551–3561. https://doi.org/10.1519/JSC.0000000000004142 (2022).

39. Cohen, J. A power primer. *Psychol. Bull.* **112**(1), 155–159 (1992).

40. World Medical Association. World medical association declaration of helsinki: Ethical principles for medical research involving human subjects. *JAMA* **310**(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053 (2013).

41. Nilstun, T. Nya Forskningsetiska Riktlinjer Från CIOMS. Föredömlig avvägning autonomi-nytta-rättvisa. *Lakartidningen* **91**(3), 157–161 (1994).

42. Seaw, K. M., Leow, M. K. S. & Bi, X. Early obesity risk prediction via non-dietary lifestyle factors using machine learning approaches. *Clin. Obes.* **15**(1), e70011. https://doi.org/10.1111/cob.70011 (2025).

43. Syahidah, H., Irsandi, N., Nur Ajizah, A. & Amelia, A. Obesity prediction using machine learning algorithms. *Int. J. Adv. Technol. Innov. Sci.* **2**(1), 1 (2025). https://journal.irpi.or.id/index.php/ijatis

44. Dirik, M. Application of machine learning techniques for obesity prediction: A comparative study. *J. Complex. Health Sci.* **6**(2), 16–34. https://doi.org/10.21595/chs.2023.23193 (2023).

45. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794. https://doi.org/10.1145/2939672.2939785 (2016).

46. Benhar, H., Idri, A. & Fernández-Alemán, J. L. Data preprocessing for heart disease classification: A systematic literature review. *Comput. Methods Programs Biomed.* **195**, 105635. https://doi.org/10.1016/j.cmpb.2020.105635 (2020).

47. Wu, Y., Li, D. & Vermund, S. H. Advantages and limitations of the body mass index (BMI) to assess adult obesity. *Int. J. Environ. Res. Public Health* **21**(6), 757. https://doi.org/10.3390/ijerph21060757 (2024).

48. Bosch, T. A. et al. Visceral adipose tissue measured by DXA correlates with measurement by CT and is associated with cardiometabolic risk factors in children. *Pediatr. Obes.* **3**, 172–179. https://doi.org/10.1111/ijpo.249 (2015).

49. Jin, M. et al. Characteristics and reference values of fat mass index and fat free mass index by bioelectrical impedance analysis in an adult population. *Clin. Nutr.* **38**(5), 2325–2332. https://doi.org/10.1016/j.clnu.2018.10.010 (2019).

50. Peltz, G., Aguirre, M. T., Sanderson, M. & Fadden, M. K. The role of fat mass index in determining obesity. *Am. J. Hum. Biol.* **22**(5), 639–647. https://doi.org/10.1002/ajhb.21056 (2010).

51. Liu, P., Ma, F., Lou, H. & Liu, Y. The utility of fat mass index vs. body mass index and percentage of body fat in the screening of metabolic syndrome. *BMC Public Health* **13**, 629. https://doi.org/10.1186/1471-2458-13-629 (2013).

52. Kuk, J. L. et al. Visceral fat is an independent predictor of all-cause mortality in men. *Obesity* **14**(2), 336–341. https://doi.org/10.1038/oby.2005.45 (2005).

53. Ramírez-Vélez, R. et al. Percentage of body fat and fat mass index as a screening tool for metabolic syndrome prediction in Colombian university students. *Nutrients* **9**(9), 1009. https://doi.org/10.3390/nu9091009 (2017).

54. Huang, A. A. & Huang, S. Y. Increasing transparency in machine learning through bootstrap simulation and shapely additive explanations. *PLoS One* **18** (2), e0281922. https://doi.org/10.1371/journal.pone.0281922 (2023).

55. Zhou, Y. et al. Distinguishing apathy and depression in older adults with mild cognitive impairment using text, audio, and video based on multiclass classification and shapely additive explanations. *Int. J. Geriatr. Psychiatry.* https://doi.org/10.1002/gps.5827 (2022).

56. Lundberg, S. M. & Lee, S. I. A Unified approach to interpreting model predictions. *Adv. Neural Inf. Process Syst.* **30**, 4765–4774 (2017).

57. Lin, W., Shi, S., Huang, H., Wen, J. & Chen, G. Predicting risk of obesity in overweight adults using interpretable machine learning algorithms. *Front. Endocrinol.* **14**, 1292167. https://doi.org/10.3389/fendo.2023.1292167 (2023).

58. Barber, T. M., Kabisch, S., Pfeiffer, A. F. & Weickert, M. O. Optimised skeletal muscle mass as a key strategy for obesity management. *Metabolites* **15**(2), 85. https://doi.org/10.3390/metabo15020085 (2025).

59. AlMasud, A. A. et al. Relationship of fat mass index and fat free mass index with body mass index and association with sleeping patterns and physical activity in Saudi young adults women. *J. Health Popul. Nutr.* **44**(1), 64. https://doi.org/10.1186/s41043-025-00795-5 (2025).

60. Butte, N. F. et al. Energetic adaptations persist after bariatric surgery in severely obese adolescents. *Obesity* **23**(3), 591–601. https://doi.org/10.1002/oby.20994 (2015).

61. Yang, R. et al. Correlations and consistency of body composition measurement indicators and BMI: A systematic review. *Int. J. Obes.* **49**(1), 4–12. https://doi.org/10.1038/s41366-024-01638-9 (2025).

62. Bosy-Westphal, A. & Müller, M. J. Diagnosis of obesity based on body composition-associated health risks—Time for a change in paradigm. *Obes. Rev.* **22**(2), e13190. https://doi.org/10.1111/obr.13190 (2021).

63. Abraham, A. & Yaghootkar, H. Identifying obesity subtypes: A review of studies utilising clinical biomarkers and genetic data. *Diabet. Med.* **40**(12), e15226. https://doi.org/10.1111/dme.15226 (2023).

Acknowledgements

Author contributions

R.Y.-S. conceptualized the study, designed the research, conducted data analysis, drafted the manuscript, and coordinated the research team. A.V.-B. was responsible for data curation, statistical analysis, and contributed to manuscript writing. R.O. supported the literature review, data analysis, and preparation of figures and tables. P.O. participated in data collection and field coordination, as well as in result interpretation. J.P.Z.-C. contributed to methodology development and manuscript editing. C.H.-T. collaborated in data collection and quality control. C.M.-S. assisted in literature search and manuscript formatting. F.G.-R. contributed to the development of analytical models and critical review of the findings. J.d.S.-L. assisted in data interpretation and translation of technical content. J.P.-H. participated in data verification and graphical visualization. J.O.-A. contributed to the

Declarations

Competing interests
The authors declare no competing interests.

Consent for publication
All authors have agreed to the publication of this manuscript.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-15264-6.

Correspondence and requests for materials should be addressed to J.F.L.-G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.