Prevention of Cardiometabolic Syndrome in Children and Adolescents Using Machine Learning and Noninvasive Factors: The CASPIAN-V Study

Hamid Reza Marateb [1,*], Mahsa Mansourian [2], Amirhossein Koochekian [3], Mehdi Shirzadi [1], Shadi Zamani [4], Marjan Mansourian [1,*], Miquel Angel Mañanas [1,5] and Roya Kelishadi [3]

1   Biomedical Engineering Research Centre (CREB), Automatic Control Department (ESAII), Universitat Politècnica de Catalunya-Barcelona Tech (UPC), 08028 Barcelona, Spain; mehdi.shirzadi@upc.edu (M.S.); miguel.angel.mananas@upc.edu (M.A.M.)
2   Department of Medical Physics, School of Medicine, Isfahan University of Medical Sciences, Isfahan 81746-73461, Iran; mansourian@med.mui.ac.ir
3   Child Growth and Development Research Center, Research Institute for Primordial Prevention of Non-Communicable Disease, Isfahan University of Medical Sciences, Isfahan 81746-73461, Iran; kelishadi@med.mui.ac.ir (R.K.)
4   Biomedical Engineering Department, Engineering Faculty, University of Isfahan, Isfahan 81746-73441, Iran
5   Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), 28029 Madrid, Spain
*   Correspondence: hamid.reza.marateb@upc.edu (H.R.M.); marjan.mansourian@upc.edu (M.M.); Tel.: +34-934-01-16-93 (H.R.M.)

Abstract: Cardiometabolic syndrome (CMS) is a growing concern in children and adolescents, marked by obesity, hypertension, insulin resistance, and dyslipidemia. This study aimed to predict CMS using machine learning based on data from the CASPIAN-V study, which involved 14,226 participants aged 7–18 years, with a CMS prevalence of 82.9%. We applied the XGBoost algorithm to analyze key noninvasive variables, including self-rated health, sunlight exposure, screen time, consanguinity, healthy and unhealthy dietary habits, discretionary salt and sugar consumption, birthweight, and birth order, father and mother education, oral hygiene behavior, and family history of dyslipidemia, obesity, hypertension, and diabetes using five-fold cross-validation. The model achieved high sensitivity (94.7% ± 4.8) and specificity (78.8% ± 13.7), with an area under the ROC curve (AUC) of 0.867 ± 0.087, indicating strong predictive performance and significantly outperformed triponderal mass index (TMI) (adjusted paired *t*-test; $p < 0.05$). The most critical selected modifiable factors were sunlight exposure, screen time, consanguinity, healthy and unhealthy diet, dietary fat type, and discretionary salt consumption. This study emphasizes the clinical importance of early identification of at-risk individuals to implement timely interventions. It offers a promising tool for CMS risk screening. These findings support using predictive analytics in clinical settings to address the rising CMS epidemic in children and adolescents.

Keywords: adolescent health; cardiometabolic risk; children; early diagnosis; machine learning; metabolic syndrome; pediatric cardiology; predictive analytics; risk assessment; screening programs

1. Introduction

Cardiometabolic syndrome (CMS) is a common metabolic dysfunction characterized by having at least one metabolic abnormalities or risk factor, including abnormal anthropometric measures, insulin resistance, systemic arterial hypertension, and hyperlipidemia [1]. The World Health Organization (WHO), the American Society of Endocrinology, and the National Cholesterol Education Program (NCEP) have recognized CMS as a disease entity [2]. About 25% of adults worldwide suffer from CMS [3]. Recent studies have also reported a contemporaneous growth in the incidence and prevalence of youth onset

noncommunicable disease that is particularly evident among minority populations [4]. Hence, detecting metabolic risk factors in childhood is important to prevent future diabetes. Hence, identifying childhood metabolic risk factors is important for preventing future noncommunicable diseases (NCDs) [5].

Cardiovascular disease (CVD), the most common NCD, remains the leading cause of death in the Middle East, with ischemic stroke, peripheral vascular diseases, coronary heart disease, cerebrovascular arterial diseases, and myocardial infarctions [6]. Approximately 20–23% of the burden of diseases and 46% of all deaths in Iran were attributed to CVDs [7]. Also, the latest Global Burden of Disease (GBD) version revealed that Iran was one of the countries with the highest rate of CVD, with more than 9000 cases stricken by CVD per 100,000 persons [8]. The sharp rise in cases of cardiometabolic disorders, including type 2 diabetes and hypertension, highlights the urgent need for improved prevention strategies. This shift should move away from a universal approach towards a more individualized, precision-based method for the broader population [9].

Cardiometabolic syndrome in children implicates the early start of metabolic dysfunction and related metabolic risk factors (e.g., high blood pressure (BP), dyslipidemia, insulin resistance, hypertension, etc.) in childhood, which are important in the development of the risk of CVDs, and type 2 diabetes mellitus in adulthood [10]. Therefore, predicting the onset of CMS in advance is significant, which can prevent it from evolving into a more serious disease. Identifying high-risk individuals on a trajectory to develop the disease is used to focus on modifiable risk factors.

In recent years, machine learning (ML) methods have verified improved performance in predicting various diseases or clinical conditions. Compared with other traditional statistical methods like regression modeling, ML models can capture multifaceted, non-linear relationships and identify unknown correlations in big data, thus enabling deeper perceptions to be gained from clinical data [11]. With the help of such ML methods, noninvasive indirect indicators that do not require blood samples can be applied to predict CMS, enabling early diagnosis of metabolic syndrome (MS) and CMS in children and adolescents.

While machine learning methods, including the XGBoost algorithm, have been applied to predict various health conditions, the existing studies in this domain predominantly focus on MS rather than CMS [12,13]. Few studies have been performed to predict cardiometabolic risk factors (CMRs) using noninvasive and low-cost diagnosis methods. Kupusinac et al. developed an initial estimation of CMR using artificial neural networks (ANNs) to assess cardiometabolic risk in a Serbian cohort of 1281 individuals, achieving a model accuracy of 82.76% [14]. The selected features were gender, age, waist-to-height ratio, body mass index, and systolic and diastolic BP, and the ages of the participants ranged from 18 to 67 years old. Regression logistics is the other approach used by Hollenbeck et al. for predicting the prevalence of CMR factors on the National Health and Nutrition Examination Survey (NHANES) in the United States with a study sample of 5275 adults, achieving an area under the ROC curve (AUC) of 0.77 for abdominal obesity, indicating good discrimination [15]. This study included all adults aged 18 or older (an average of 45). Ashley-Martin et al. examined the predictive power of the triponderal mass index (TMI) and body mass index (BMI) for cardiometabolic markers in 5814 Canadian children aged 6–19 years, finding both indices to have similar predictive accuracy, achieving AUC values of 0.81 for insulin resistance and 0.83 for having ≥3 abnormal tests [16]. Ojanen et al. [17] conducted a comprehensive study on the early prediction of cardiometabolic syndrome in children and adolescents, following a cohort of 396 Finnish girls (mean age 11.2 years) to early adulthood. They identified key serum metabolic biomarkers such as glycoprotein acetyls (GlycA), large high-density lipoprotein phospholipids (L-HDL-PL), and the apolipoprotein B to apolipoprotein A-1 (ApoB/ApoA) ratio. Using LASSO regression and random intercept cross-lagged panel models, the study achieved AUC values between 0.641 and 0.802, explaining 32.7% to 36.3% of the variance in MS scores, with findings validated across a larger longitudinal cohort of 2664 participants and multiple

cross-sectional datasets. Among the studies mentioned above, Ashley-Martin et al. [16] and Ojanen et al. [17] focused on diagnosing CMS in children and adolescents.

Our research addresses this critical gap by specifically targeting CMS. This more comprehensive construct integrates additional cardiovascular risk factors beyond those typically considered in MS. This distinction is vital, as CMS encompasses a broader array of risk factors, such as BMI, total cholesterol (TC), and low-density lipoprotein-cholesterol (LDL-C), which are more directly linked to cardiovascular diseases. Our study applies machine learning methods, particularly extreme gradient boosting (XGBoost), to predict CMS in children using noninvasive factors, providing a more holistic and accurate assessment of cardiovascular risks from an early age.

The rationale for focusing on CMS over MS lies in its ability to offer a more comprehensive evaluation of a child's health by considering both metabolic and cardiovascular risks. This broader perspective is crucial for early identification and intervention in children who may be at an elevated risk of cardiovascular diseases later in life. By addressing CMS, our study aims to provide a robust and comprehensive tool for early risk stratification in pediatric populations. Rather than the more narrowly defined MS, this focus on CMS represents a significant and novel contribution to the field. We have thoroughly revised our manuscript to highlight this distinction, emphasizing the importance of our approach in advancing the prediction and prevention of cardiovascular diseases from an early age. Furthermore, our research stands out in the literature by being the first to explore and predict CMS using a variety of noninvasive risk factors in a pediatric population, filling a critical gap and offering a new direction for future studies.

2. Materials and Methods

2.1. Materials

Data for this study were taken from part of the "National survey of school student high risk 86 behaviors", started in 2014. The data were collected in a comprehensive database entitled "The Prevalence and Dissemination of Non-Communicable Diseases in Children and Adolescents (CASPIAN-V)" [18].

The study population was children aged 7–18 years; all were in primary and secondary schools and came from urban and rural areas. The sampling was conducted in 30 provinces of Iran according to sex (male and female), place of residence (rural and urban), and educational level (primary and secondary) via stratified clustering sampling method [18]. The dataset is a collection of 510 features with a sample size 14,226. The protocol was mainly based on the WHO-Global School student Health Survey. A minimal dataset is included in the Supplementary Material.

2.2. Procedure and Measurements

2.2.1. Questionnaires

Two sets of questionnaires were developed for both students and their parents. The student questionnaire was adapted from the Persian-translated version of the WHO's Global School-based Student Health Survey (WHO-GSHS). The reliability and validity of these questionnaires have been previously established. Eligible students were identified, and questions regarding their health status and health-related behaviors were explained using simple language. Subsequently, the students' parents were invited to complete a separate questionnaire, which collected information on family characteristics, including household size, socioeconomic variables, and the student's birth order. A team of healthcare professionals supervised and controlled the process [19]. The questionnaire included different aspects, such as variables related to students' schools, counseling with family members, and parameters related to health status, living environment, screen time, and lifestyle components.

2.2.2. Anthropometric Measurements

The information was recorded by a team of trained healthcare experts based on approved checklists. All team members were instructed to calibrate the data according to protocols and standards. The physical measurements were as follows: BMI defines obesity and is calculated as the ratio of body mass and the square of body height. The subjects were asked to wear light clothes, and their body weight was measured on a scale placed on flat ground with a precision of 0.1 kg. Their height was measured without shoes to the nearest 0.1 cm. The WHO growth charts were used to categorize BMI [20]. The waist circumferences were measured with nonelastic tape with a precision of 0.1 cm around the point halfway between the iliac crest and the sides of the lower ribs [21]. Hip circumference was measured at the widest point around the buttocks to the nearest 0.1 cm [21]. Wrist circumference was measured with the precision of 0.1 cm using a tape meter on the dominant arm while subjects held their arm on a flat surface like a table. The measuring tape was positioned over the Lister tubercle of the distal radius and the distal ulna [18]. Neck circumference was measured using nonelastic tape with an accuracy of 0.1 cm at the lower margin of the thyroid cartilage, with the head erect [18].

2.2.3. Blood Pressure Measurement

BP was measured on the right arm using a mercury sphygmomanometer with an appropriately sized cuff. BP measurement was carried out two times at 5 min intervals. Then, the average systolic and diastolic pressures were registered [22].

2.2.4. Blood Sampling

Blood samples were drawn from selected students. After 12 h overnight fasting, a 6 mL venous blood sample was collected from the students whose parents accompanied them. The blood samples were centrifuged at 2500–3000$\times$ *g* for 10 min. The resulting serum was divided into 200-microliter tubes and stored at -70 °C. Then, the samples were transferred to a referral laboratory by a cold chain. Fasting blood glucose (FBG), TC, triglyceride (TG), high-density lipoprotein-cholesterol (HDL-C), and LDL-C concentrations were determined using a Hitachi autoanalyzer (Tokyo, Japan). The flow chart of the CASPIAN-V study is presented in Figure 1.

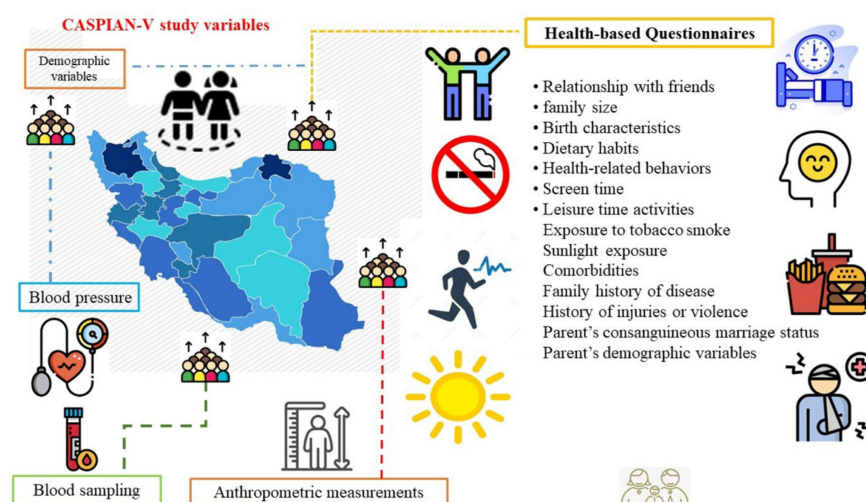**Figure 1.** Flow chart of the Prevalence and Dissemination of Non-Communicable Diseases in Children and Adolescents (CASPIAN-V) study design.

2.2.5. Outcome Definition

We used metabolic syndrome defined as central obesity (waist circumference (WC) $\geq$ 80 cm for women; WC $\geq$ 90 cm for men aged 16 years or older; WC > 90th percentile for age and sex) plus any two or more of the following components: high FBG

(FBG ≥ 100 mg/dL), high BP (systolic and/or diastolic BP > 90th percentile for age, sex, and height), low HDL-C (HDL-C < 40 mg/dL), and high TG (TG > 150 mg/dL), according to the International Diabetes Federation (IDF) [1,23]. Furthermore, BMI ≥ 85th percentile for age and sex, TC > 200 mg/dL, and LDL-C > 130 mg/dL were considered as other CMS components following the American Heart Association [24] and American Diabetes Association [25], for children and adolescents at high risk of CVD (a.k.a., cardiometabolic risks; CMR). A total of 6832 subjects with CMR labels were included in the analysis.

2.2.6. Ethical Statement

This study received approval from the Research and Ethics Council of Isfahan University of Medical Sciences (project number: 194049) and complied with the Declaration of Helsinki guidelines. Both written and verbal informed consent were obtained from parents and students following a detailed explanation of the study's goals and procedures.

2.3. Predictors and Feature Extraction

All participants were assessed according to demographic variables, anthropometric measures, laboratory variables, blood pressure, and health-related variables such as relationship with friends, family size, birth characteristics, and dietary habits.

Health-related behaviors, screen time, leisure time activities, smoking status, sunlight exposure, comorbidities, family history of the disease, student injuries, parent's consanguineous marriage, and demographic variables were recorded. The details regarding the related questions and the categorization methods for each variable are provided in Tables 1 and 2. Table 1 includes 20 input variables derived from single-item questions, while Table 2 presents four composite variables.

Table 1. Questions, response options, and references for the rationale of including single-item variables in the questionnaire.

Item	Type	Description
1	Q	Oral hygiene behavior [26]: How many times per day do you brush your teeth?
	A	(1) once daily, (2) twice daily, (3) three times a day, (4) four times and more a day, and (5) never
2	Q	The number of close friends [26]
	A	(1) none, (2) one friend, (3) two friends, and (4) three or more friends
3, 4	Q	Father's education; Mother's education
	A	(1) illiterate, (2) Quranic Literacy, (3) primary, (4) intermediate, (5) diploma, (6) Bachelor, (7) upper than bachelor, (8) died
5	Q	Number of family members
	A	Count variable
6	Q	Consanguineous marriage [27]
	A	(1) no, (2) yes
7	Q	Birth order [28]
	A	Count variable
8	Q	Birthweight category [29]
	A	(1) <2500 g, (2) 2500–400 g, (3) >4000 g
9	Q	Breastfeeding [30]: How many months did breastfeeding occur in the first two years after birth?
	A	Count variable
10–14	Q	Family history of hypertension [31]; Family history of dyslipidemia [32]; Family history of diabetes [33]; Family history of obesity [34]; Family history of cancer
	A	(1) no, (2) yes

Table 1. *Cont.*

Item	Type	Description
15	Q	Residence [34]
	A	(1) urban, (2) rural
16	Q	Age group
	A	7–10, B. 11–14- and C. 15–18-year-old
17	Q	Gender [35]
	A	(1) male, (2) female
18	Q	Self-rated health [36]: How would you describe your general state of health?
	A	(1) perfect, (2) good, (3) bad, and (4) very bad
19	Q	Discretionary salt
	A	(1) always, (2) sometimes, (3) rarely, and (4) no
20	Q	Dietary fat type [37]
	A	Saturated fats, (2) trans fats, (3) monounsaturated fats, and (4) polyunsaturated fats.

Q: question; A: answer.

Table 2. Questions, response options, and combination methods for composite variables in the questionnaire.

Item	Type	Description
21	V	Healthy diet
	Q	Five questions: Diet beverage, fresh fruit, dried fruit, fresh fruit juice, and fresh/boiled vegetable consumption
	A	"Daily", "weekly", "rarely", and "never"
	C	The answers were summed up, and the tertiles were calculated, resulting in a three-category ordinal variable ("low", "moderate", and "high").
22	V	Unhealthy diet
	Q	Six questions: Sweets, fast food, soda, junk food, sugar-sweetened fruit juice, and discretionary sugar consumption
	A	"Daily," "weekly", "rarely", and "never"
	C	The answers were summed up, and the tertiles were calculated, resulting in a three-category ordinal variable ("low", "moderate", and "high").
23	V	Screen time
	Q	Six questions: Q1, Q2: How many hours a day do you watch TV (on weekdays/weekends)? Q3, Q4: How many hours a day do you do your class exercises (on weekdays/weekends)? Q5, Q6: How many hours a day do you use a computer (on weekdays/weekends)?
	A	(1) never, (2) 1 h, (3) 2 h, (4) 3 h, and (5) four or more hours.
	C	The answers to the questions were summed up, and the cut-off of 10 was used to create a dichotomous variable (<10: No, ≥10: Yes). The optimal" cut-point was calculated to minimize the error rate (ER) criteria [38] in CMR ROC Analysis.

Table 2. *Cont.*

Item	Type	Description
	V	Sunlight exposure [39]
24	Q	Four questions: Q1, Q2: How much exposure to outdoor sunlight (on weekdays/weekends)? Q3: What parts of the body are exposed to sunlight during sunlight exposure? Q4: Do you use sunscreen creams?
	A	Q1, Q2: (1) less than 5 min, (2) between 5 and 30 min, and (3) more than 30 min Q3: (1) hands, face, legs, (2) hands, face, arms, and (3) hands, face Q4: (1) never, (2) sometimes, and (3) always
	C	Using Principal Component Analysis (PCA) [11], the primary principal component was extracted from the responses to the questions. Subjects were subsequently classified into "low", "medium", and "high" sunlight exposure groups according to the tertiles of the principal component.

V: variable; Q: questions; A: answer; C: composite variable generation method.

2.4. Model Construction and Interpretation

The categorical variables were transferred into an interval scale using the logit function (Equation (1)), whose parameters were estimated by maximum likelihood estimating (MLE) on the training set [40]:

$$p(I(x)) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \times I(x))}} \tag{1}$$

where β_0 and β_1 are the intercept and rate parameters, respectively, and $I(x)$ is an indicator variable of the categorical input variable. An indicator variable is an axillary binary variable showing the category of a categorical variable. For example, the "unhealthy diet" categorical variable has three groups ("low", "medium", and "high"). It, thus, has two indicator variables of "medium" and "high". The reference "low" category is not used, as it creates multicollinearity in the predictors [41].

Since there was some information missing in some of the variables, the missing completely at random (MCAR) assumption was examined using Little's MCAR test ($p = 0.106$) [42]. Multivariate imputation by chained equations (MICE) [43] was then used for missing data imputation.

For classifying subjects with CMR from healthy subjects, we used the XGBoost algorithm [44]. XGBoost is a scalable machine learning algorithm for tree boosting that can solve complex problems with different sizes of datasets. It is a free library that efficiently implements a gradient-boosting framework [45,46].

The XGBoost algorithm uses n samples and m features as a given data to predict the target with a tree ensemble model with K different trees:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \; f_k \in F \tag{2}$$

where $F = \{f(x) = wq(x)\}f$, and $q : R^m \rightarrow T$, $w \in R^T$ is the space of the regression tree. In this formula, q is related to the structure of each tree, and it can map a sample to a leaf. T is the number of leaves in the tree and f_k is an independent tree with a structure of q and a leaf with weight of w. This function gives each sample to different trees, and, at the end, it sums all of the results. Because we have a regression tree, the scores on each leaf are continuous. For each sample, classification is performed by applying the decision rules from the trees, leading to a final prediction based on the sum of all scores in the leaves. The model learns the set of functions by minimizing the following regularization function:

$$L(\phi) = \sum_i l(\hat{y}_i, \; y_i) + \sum_k \Omega(f_k) \tag{3}$$

where $\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2$ and l represent the differentiable convex loss function that quantifies the discrepancy between the prediction and the target. The second term acts as a penalty for model complexity to prevent overfitting.

To tune the hyperparameters of the XGBoost algorithm, we used grid search (brute force) [47,48]. After hyperparameter optimization, we separated the dataset into five subsets using stratified random sampling for five-fold cross-validation.

2.5. Validation

Five-fold cross-validation was used in our study to guard against testing hypotheses suggested by the data (Type III errors). The signal detection theory parameters (true positive (TP), true negative (TN), false positive (FP), and false negative (FN)) were then calculated on each test fold. The performance indices introduced as the following were calculated. Finally, the indices' mean and standard deviation were provided over the test folds, and the cross-validated indices and their confidence interval (CI) of 95% were reported, following the Standards for Reporting Diagnostic Accuracy (STARD) and Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines [49].

TPs are cases where the model correctly predicts CMS, while FPs are incorrect predictions of CMS in patients who do not have it. FNs occur when the model fails to identify CMS in affected patients, and TNs are correct predictions of no CMS in unaffected patients. These outcomes are essential for evaluating the model's performance:

$$Se = \frac{TP}{TP + FN} \tag{4}$$

$$Sp = \frac{TN}{TN + FP} \tag{5}$$

$$PPV = \frac{TP}{TP + FP} \tag{6}$$

$$NPV = \frac{TN}{TN + FN} \tag{7}$$

where *Se*, *Sp*, *PPV*, and *NPV* are sensitivity, specificity, positive predictive value, and negative predictive value, respectively.

We further calculated the following composite performance indices based on the ISO/IEC TS:2022 standard [50]:

$$LR^+ = \frac{Se}{1 - Sp} \tag{8}$$

$$LR^- = \frac{1 - Se}{Sp} \tag{9}$$

$$DOR = \frac{TP \times TN}{FP \times FN} \tag{10}$$

$$DP = \left(\sqrt{\frac{3}{\pi}}\right) \times log(DOR) \tag{11}$$

$$AUC = \frac{Se + Sp}{2} \tag{12}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{13}$$

$$Kappa = \frac{2 \times (TP \times TN - FP \times FN)}{(TP + FP)(TP + FN) \times (TN + FP) \times (TN + FN)} \tag{14}$$

$$F_1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{15}$$

where LR^+, LR^-, DOR, DP, AUC, MCC, *Kappa*, and F_1 are positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, discriminant power, area under the receiver operating characteristic (ROC) curve (balanced accuracy), Matthew's correlation coefficient, Cohen's kappa coefficient (also known as the agreement between the predicted class labels and the gold standard), and F_1 score, respectively.

2.6. Statistical Analysis

Results were reported frequencies (for categorical variables) and mean $\pm$ standard deviation (for interval variables). *p*-values less than 0.05 were considered significant. The association between the interval variables with normal distribution to the CMR or healthy group was tested using the independent-sample *t*-test. Mann–Whitney U-test was used for interval variables with non-normal distribution or ordinal data. For the entire nominal variables, the χ^2 test was used since frequencies in all groups were greater than five. Fisher's exact test was used otherwise. We compared the performance of our proposed method with the TMI index using an adjusted paired *t*-test [51]. The statistical analysis and calculations were performed using the SPSS statistical package, version 22.0 (IBM Corp. Released 2013. Armonk, NY, USA: IBM Corp.).

3. Results
3.1. Descriptive and Inferential Statistics

The average age of the participants was 12.4 $\pm$ 3.0, and 48.6% were female. A total of 5663 subjects (82.9%) were in the CMR group. The characteristics of the participants in the healthy and CMR group are shown in Table 3.

Table 3. Demographic and behavioral characteristics of participants by cardiometabolic syndrome status.

Characteristics	Non-CMS (N = 1169)	CMS (N = 5663)	Overall (N = 6832)	*p*-Value
Number of close friends				
None	54 (4.6%)	240 (4.2%)	294 (4.3%)	0.3932
One friend	212 (18.1%)	934 (16.5%)	1146 (16.8%)	
Two friends	333 (28.5%)	1594 (28.1%)	1927 (28.2%)	
Three or more friends	570 (48.8%)	2895 (51.1%)	3465 (50.7%)	
Oral hygiene behavior				
Never	47 (4.0%)	225 (4.0%)	272 (4.0%)	0.6016
Less than once a week	52 (4.4%)	204 (3.6%)	256 (3.7%)	
At least once a week	124 (10.6%)	666 (11.8%)	790 (11.6%)	
Only once a week	62 (5.3%)	289 (5.1%)	351 (5.1%)	
Once a day	626 (53.6%)	2974 (52.5%)	3600 (52.7%)	
More than once a day	258 (22.1%)	1305 (23.0%)	1563 (22.9%)	
Self-rated health				
Average	241 (20.6%)	1023 (18.1%)	1264 (18.5%)	0.0640
Bad	4 (0.3%)	46 (0.8%)	50 (0.7%)	
Good	428 (36.6%)	2075 (36.6%)	2503 (36.6%)	
Perfect	496 (42.4%)	2519 (44.5%)	3015 (44.2%)	
Age group				
7–10 y	396 (33.9%)	1737 (30.7%)	2133 (31.2%)	0.0107
11–14 y	505 (43.2%)	2404 (42.5%)	2909 (42.6%)	
15–18 y	268 (22.9%)	1522 (26.9%)	1790 (26.2%)	
Residence				
Rural	360 (30.8%)	1366 (24.1%)	1726 (25.3%)	<0.001
Urban	809 (69.2%)	4297 (75.9%)	5106 (74.7%)	
Number of family members				
Mean (SD)	4.91 (1.54)	4.81 (1.47)	4.83 (1.48)	0.0631
Median [Min, Max]	5.00 [1.00, 17.0]	5.00 [0, 15.0]	5.00 [0, 17.0]	

Characteristics	Non-CMS (N = 1169)	CMS (N = 5663)	Overall (N = 6832)	*p*-Value
Father education				
Upper than bachelor	53 (4.6%)	280 (5.0%)	333 (4.9%)	0.0413
Bachelor	104 (8.9%)	586 (10.3%)	690 (10.1%)	
Diploma	264 (22.6%)	1398 (24.7%)	1662 (24.3%)	
Intermediate	290 (24.8%)	1366 (24.1%)	1656 (24.2%)	
Primary	264 (22.6%)	1240 (21.9%)	1504 (22.0%)	
Quranic Literacy	17 (1.5%)	55 (1.0%)	72 (1.1%)	
Illiterate	158 (13.5%)	612 (10.8%)	770 (11.3%)	
Father died	19 (1.6%)	126 (2.2%)	145 (2.1%)	
Mother education				
Upper than bachelor	15 (1.3%)	91 (1.6%)	106 (1.6%)	0.3919
Bachelor	106 (9.1%)	577 (10.2%)	683 (10.0%)	
Diploma	291 (24.9%)	1504 (26.6%)	1795 (26.3%)	
Intermediate	254 (21.7%)	1135 (20.0%)	1389 (20.3%)	
Primary	281 (24.0%)	1398 (24.7%)	1679 (24.6%)	
Quranic Literacy	13 (1.1%)	57 (1.0%)	70 (1.0%)	
Illiterate	205 (17.5%)	876 (15.5%)	1081 (15.8%)	
Mother died	4 (0.3%)	25 (0.4%)	29 (0.4%)	
Consanguinity				
No	651 (55.7%)	3068 (54.2%)	3719 (54.4%)	0.3612
Yes	518 (44.3%)	2595 (45.8%)	3113 (45.6%)	
Birth order				
Mean (SD)	2.27 (1.59)	2.17 (1.49)	2.19 (1.51)	0.0485
Median [Min, Max]	2.00 [0, 11.0]	2.00 [0, 11.0]	2.00 [0, 11.0]	
Birthweight category				
2500–4000 g	859 (73.5%)	4233 (74.7%)	5092 (74.5%)	0.2791
Less than 2500 g	122 (10.4%)	507 (9.0%)	629 (9.2%)	
More than 4000 g	188(16.1%)	923 (16.3%)	1111 (16.3%)	
Breastfeeding				
Mean (SD)	15.0 (8.52)	15.1 (8.52)	15.1 (8.52)	0.735
Median [Min, Max]	18.0 [0, 24.0]	18.0 [0, 24.0]	18.0 [0, 24.0]	
Discretionary salt				
Always	237 (20.3%)	1268 (22.4%)	1505 (22.0%)	0.0154
Sometimes	255 (21.8%)	1131 (20.0%)	1386 (20.3%)	
Rarely	116 (9.9%)	705 (12.4%)	821 (12.0%)	
No	561 (48.0%)	2559 (45.2%)	3120 (45.7%)	
Family history of hypertension				
No	452 (38.7%)	2178 (38.5%)	2630 (38.5%)	0.9216
Yes	717 (61.3%)	3485 (61.5%)	4202 (61.5%)	
Family history of dyslipidemia				
No	587 (50.2%)	2846 (50.3%)	3433 (50.2%)	1.0000
Yes	582 (49.8%)	2817 (49.7%)	3399 (49.8%)	
Family history of diabetes				
No	552 (47.2%)	2752 (48.6%)	3304 (48.4%)	0.4093
Yes	617 (52.8%)	2911 (51.4%)	3528 (51.6%)	
Family history of obesity				
No	635 (54.3%)	2879 (50.8%)	3514 (51.4%)	0.0327
Yes	534 (45.7%)	2784 (49.2%)	3318 (48.6%)	
Family history of cancer				
No	944 (80.8%)	4552 (80.4%)	5496 (80.4%)	0.8018
Yes	225 (19.2%)	1111 (19.6%)	1336 (19.6%)	
Dietary fat type				
Monounsaturated fats	478 (40.9%)	2109 (37.2%)	2587 (37.8%)	0.0362
Polyunsaturated fats	181 (15.5%)	1041 (18.4%)	1222 (17.9%)	
Saturated fats	429 (36.7%)	2083 (36.8%)	2512 (36.8%)	
Trans fats	81 (6.9%)	430 (7.6%)	511 (7.5%)	
Gender				
Female	602 (51.5%)	2719 (48.0%)	3321 (48.6%)	0.0326
Male	567 (48.5%)	2944 (52.0%)	3511 (51.4%)	

Characteristics	Non-CMS (N = 1169)	CMS (N = 5663)	Overall (N = 6832)	*p*-Value
Diet beverage consumption				
Daily	7 (0.6%)	46 (0.8%)	53 (0.8%)	0.2754
Weekly	49 (4.2%)	275 (4.9%)	324 (4.7%)	
Rarely	210 (18.0%)	1116 (19.7%)	1326 (19.4%)	
Never	903 (77.2%)	4226 (74.6%)	5129 (75.1%)	
Fresh fruit consumption				
Daily	666 (57.0%)	2928 (51.7%)	3594 (52.6%)	0.0101
Weekly	386 (33.0%)	2118 (37.4%)	2504 (37.6%)	
Rarely	80 (6.8%)	441 (7.8%)	521 (7.6%)	
Never	37 (3.2%)	176 (3.1%)	213 (3.1%)	
Dried fruit consumption				
Daily	400 (34.2%)	1804 (31.9%)	2204 (32.3%)	0.0354
Weekly	477 (40.8%)	2410 (42.6%)	2887 (42.3%)	
Rarely	209 (17.9%)	1134 (20.0%)	1343 (19.7%)	
Never	83 (7.1%)	315 (5.6%)	398 (5.8%)	
Fresh fruit juice consumption				
Daily	196 (16.8%)	971 (17.1%)	1167 (17.1%)	0.1903
Weekly	390 (33.3%)	2001 (35.3%)	2391 (35.0%)	
Rarely	478 (40.9%)	2276 (40.2%)	2754 (40.3%)	
Never	105 (9.0%)	415 (7.3%)	520 (7.6%)	
Fresh or boiled vegetable consumption				
Daily	347 (29.7%)	1899 (33.5%)	2246 (32.9%)	0.0511
Weekly	595 (50.9%)	2652 (46.8%)	3247 (47.5%)	
Rarely	158 (13.5%)	774 (13.7%)	932 (13.6%)	
Never	69 (5.9%)	338 (6.0%)	407 (6.0%)	
Sweets consumption				
Daily	286 (24.5%)	1409 (24.9%)	1695 (24.8%)	0.0143
Weekly	520 (44.5%)	2597 (45.8%)	3117 (45.6%)	
Rarely	352 (30.1%)	1537 (27.1%)	1889 (27.6%)	
Never	11 (0.9%)	120 (2.1%)	131 (1.9%)	
Fast food consumption				
Daily	61 (5.2%)	439 (7.8%)	500 (7.3%)	0.0014
Weekly	355 (30.4%)	1510 (26.7%)	1865 (27.3%)	
Rarely	600 (51.3%)	3042 (53.7%)	3642 (53.3%)	
Never	153 (13.1%)	672 (11.9%)	825 (12.1%)	
Soda consumption				
Daily	23 (2.0%)	212 (3.7%)	235 (3.4%)	0.0105
Weekly	276 (23.6%)	1237 (21.8%)	1513 (22.1%)	
Rarely	627 (53.6%)	3106 (54.8%)	3733 (54.6%)	
Never	243 (20.8%)	1108 (19.6%)	1351 (19.8%)	
Junk food consumption				
Daily	16 (1.4%)	126 (2.2%)	142 (2.1%)	0.0273
Weekly	179 (15.3%)	908 (16.0%)	1087 (15.9%)	
Rarely	718 (61.4%)	3568 (63.0%)	4286 (62.7%)	
Never	256 (21.9%)	1061 (18.7%)	1317 (19.3%)	
Sugar sweetened fruit juice consumption				
Daily	72 (6.2%)	373 (6.6%)	445 (6.5%)	0.2753
Weekly	246 (21.0%)	1306 (23.1%)	1552 (22.7%)	
Rarely	665 (56.9%)	3049 (53.8%)	3714 (54.4%)	
Never	186 (15.9%)	935 (16.5%)	1121 (16.4%)	
Discretionary sugar consumption				
Daily	566 (48.4%)	2766 (48.8%)	3332 (48.8%)	0.1030
Weekly	240 (20.5%)	1195 (21.1%)	1435 (21.0%)	
Rarely	254 (21.7%)	1080 (19.1%)	1334 (19.5%)	
Never	109 (9.3%)	622 (11.0%)	731 (10.7%)	

Characteristics	Non-CMS (N = 1169)	CMS (N = 5663)	Overall (N = 6832)	*p*-Value
Screen time Q1				
Never	46 (3.9%)	242 (4.3%)	288 (4.2%)	0.7006
1 h	467 (39.9%)	2278 (40.2%)	2745 (40.2%)	
2 h	355 (30.4%)	1626 (28.7%)	1981 (29.0%)	
3 h	200 (17.1%)	972 (17.2%)	1172 (17.2%)	
4 h or more	101 (8.6%)	545 (9.6%)	646 (9.5%)	
Screen time Q2				
Never	32 (2.8%)	190 (3.4%)	222 (3.3%)	0.4569
1 h	239 (20.4%)	1190 (21.0%)	1429 (20.9%)	
2 h	382 (32.7%)	1709 (30.2%)	2091 (30.6%)	
3 h	313 (26.8%)	1566 (27.7%)	1879 (27.5%)	
4 h or more	203 (17.4%)	1008 (17.8%)	1211 (17.7%)	
Screen time Q3				
Never	4 (0.4%)	22 (0.4%)	26 (0.4%)	0.8998
1 h	395 (33.8%)	1933 (34.1%)	2328 (34.1%)	
2 h	289 (24.7%)	1434 (25.3%)	1723 (25.2%)	
3 h	216 (18.5%)	1065 (18.8%)	1281 (18.8%)	
4 h or more	265 (22.7%)	1209 (21.3%)	1474 (21.6%)	
Screen time Q4				
Never	37 (3.2%)	221 (3.9%)	258 (3.8%)	0.2376
1 h	218 (18.6%)	1046 (18.5%)	1264 (18.5%)	
2 h	322 (27.5%)	1681 (29.7%)	2003 (29.3%)	
3 h	335 (28.7%)	1472 (26.0%)	1807 (26.4%)	
4 h or more	257 (22.0%)	1243 (21.9%)	1500 (22.0%)	
Screen time Q5				
Never	765 (65.4%)	3513 (62.0%)	4278 (62.6%)	0.1565
1 h	307 (26.3%)	1571 (27.7%)	1878 (27.5%)	
2 h	70 (6.0%)	407 (7.2%)	477 (7.0%)	
3 h	14 (1.2%)	100 (1.8%)	114 (1.7%)	
4 h or more	13 (1.1%)	72 (1.3%)	85 (1.2%)	
Screen time Q6				
Never	604 (51.6%)	2728 (48.2%)	3332 (48.7%)	0.0789
1 h	314 (26.9%)	1685 (29.8%)	1999 (29.3%)	
2 h	149 (12.7%)	690 (12.2%)	839 (12.3%)	
3 h	71 (6.1%)	352 (6.2%)	423 (6.2%)	
4 h or more	31 (2.7%)	208 (3.7%)	239 (3.5%)	
Sunlight exposure Q1				
Less than 5 min	167 (14.3%)	898 (15.9%)	1065 (15.6%)	0.0959
5–30 min	534 (45.7%)	2400 (42.4%)	2934 (42.9%)	
More than 30 min	468 (40.0%)	2365 (41.8%)	2833 (41.4%)	
Sunlight exposure Q2				
Less than 5 min	214 (18.3%)	1039 (18.3%)	1253 (18.3%)	0.4220
5–30 min	315 (26.9%)	1628 (28.7%)	1943 (28.4%)	
More than 30 min	640 (54.7%)	2996 (52.9%)	3636 (53.2%)	
Sunlight exposure Q3				
Hands, face, legs	231 (19.8%)	1034 (18.3%)	1265 (18.5%)	0.3374
Hands, face	864 (73.9%)	4224 (74.6%)	5088 (74.5%)	
Hands, face, arms	74 (6.3%)	405 (7.1%)	479 (7.0%)	
Sunlight exposure Q4				
Never	486 (41.6%)	2225 (39.3%)	2711 (39.7%)	0.1542
Sometimes	530 (45.4%)	2588 (45.7%)	3130 (45.7%)	
Always	153 (13.1%)	850 (15.0%)	1003 (14.7%)	

The descriptive and univariate statistics table highlights several key characteristics and behaviors among individuals with CMS compared to those without it. The data reveal significant demographic, social, and health-related differences.

3.1.1. Demographic Characteristics

There are statistically significant differences ($p = 0.0107$) across age groups, with a higher proportion of individuals aged 15–18 years in the CMS group (26.9%) compared to the non-CMS group (22.9%). A notable difference in residence is observed ($p < 0.001$), with a higher percentage of urban dwellers in the CMS group (75.9%) compared to the non-CMS group (69.2%).

3.1.2. Social and Family Characteristics

Differences in father's education levels are significant ($p = 0.0413$), with the CMS group having a slightly higher percentage of fathers with education beyond a bachelor's degree (5.0%) compared to the non-CMS group (4.6%). The mean birth order is slightly lower in the CMS group (2.17) compared to the non-CMS group (2.27), with a significant p-value of 0.0485.

3.1.3. Health Behaviors

No significant differences were found in oral hygiene behaviors between the groups ($p = 0.6016$), with similar proportions practicing daily oral hygiene. There is a significant difference in discretionary salt usage ($p = 0.0154$), with the CMS group reporting slightly higher daily usage (22.4% vs. 20.3%). Daily fast-food consumption is higher in the CMS group (7.8%) compared to the non-CMS group (5.2%), with a significant p-value of 0.0014.

3.1.4. Dietary Habits

The CMS group consumes less fresh fruit daily (51.7%) than the non-CMS group (57.0%), with a significant p-value of 0.0101. A significant difference ($p = 0.0354$) is observed in dried fruit consumption, with the CMS group having lower daily consumption (31.9% vs. 34.2%). There are significant differences in sweets consumption ($p = 0.0143$), with the CMS group consuming sweets more frequently daily (24.9%) compared to the non-CMS group (24.5%).

3.1.5. Family Health History and Lifestyle Factors

A significant difference is observed in the family history of obesity ($p = 0.0327$), with the CMS group having a higher percentage of individuals with a family history of obesity (49.2% vs. 45.7%). Differences in sunlight exposure are observed, although not statistically significant overall. However, the CMS group reports slightly higher exposure.

Overall, the findings suggest that individuals with cardiometabolic syndrome tend to have distinct demographic profiles, dietary habits, and health behaviors compared to those without the syndrome. These differences highlight the importance of targeted interventions considering these factors to manage and prevent CMS effectively.

3.2. Classification Results

Running XGBoost, the following features were selected: Self-rated health, sunlight exposure, screen time, consanguinity, unhealthy diet, gender, healthy diet, rural or urban residence, dietary fat type, birth order, discretionary salt, family history of hypertension, dyslipidemia, diabetes, and obesity and oral hygiene behavior.

The performance of the proposed system and the TMI index (Ashley-Martin et al. [16]) on the test sets during cross-validation and the cross-validated confusion matrix are provided in Table 4. The comparison between their ROC curves is provided in Figure 2. The cross-validated TMI cut-off was 13.50. The proposed system significantly outperformed TMI (adjusted paired t-test; $p < 0.05$).

Table 4. Performance comparison between the proposed system and the TMI index. The first row of each metric shows the results for the proposed system, while the second row presents the results for the TMI index.

Metric	Test Folds							Cross-Validated Results		
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	SD	Value	95% CI Lower	95% CI Upper
Sensitivity	0.867	0.967	0.972	0.976	0.958	0.947	0.048	0.947	0.941	0.953
	0.554	0.541	0.549	0.552	0.537	0.547	0.007	0.547	0.535	0.558
Specificity	0.600	0.700	0.845	0.854	0.940	0.788	0.137	0.787	0.764	0.811
	0.897	0.915	0.872	0.910	0.876	0.894	0.017	0.894	0.887	0.901
PPV	0.910	0.939	0.968	0.970	0.987	0.955	0.030	0.955	0.950	0.961
	0.963	0.968	0.954	0.967	0.954	0.961	0.006	0.961	0.957	0.966
NPV	0.480	0.815	0.864	0.881	0.823	0.772	0.167	0.757	0.733	0.781
	0.294	0.292	0.286	0.296	0.282	0.290	0.005	0.290	0.279	0.301
LR+	2.150	3.218	6.292	6.687	15.939	6.857	5.437	4.452	3.986	4.973
	5.400	6.330	4.280	6.130	4.330	5.290	0.864	5.150	4.530	6.060
LR-	0.230	0.048	0.033	0.028	0.045	0.077	0.086	0.067	0.060	0.076
	0.497	0.502	0.518	0.492	0.529	0.507	0.014	0.507	0.496	0.520
Kappa	0.420	0.706	0.825	0.840	0.850	0.728	0.182	0.723	0.700	0.745
	0.249	0.247	0.232	0.253	0.225	0.241	0.011	0.241	0.231	0.250
DOR	9.350	67.717	191.000	235.202	354.461	171.545	136.910	66.102	55.057	79.362
	10.87	12.600	8.270	12.450	8.190	10.480	1.930	10.16	8.760	12.200
DP	0.950	1.789	2.229	2.317	2.491	1.955	0.619	1.779	1.701	1.856
	1.013	1.075	0.897	1.070	0.893	0.990	0.081	0.984	0.918	1.061
AUC	0.730	0.833	0.909	0.915	0.949	0.867	0.087	0.867	0.858	0.876
	0.774	0.748	0.730	0.782	0.733	0.753	0.021	0.754	0.743	0.764
MCC	0.420	0.709	0.825	0.840	0.853	0.730	0.181	0.723	0.711	0.734
	0.341	0.344	0.318	0.349	0.312	0.333	0.015	0.333	0.320	0.345

CV: cross-validated; CI: confidence interval; PPV: positive predictive value; NPV: negative predictive value; LR+: positive likelihood ratio; LR-: negative likelihood ratio; DOR: diagnostics odds ratio; DP: discriminant power; AUC: area under the ROC curve; MCC: Matthew's correlation coefficient; TMI: triponderal mass index.

The classification system for cardiometabolic syndrome showed robust performance across multiple test folds, demonstrating high sensitivity and specificity values. Sensitivity averaged 0.947 (SD = 0.048), indicating that the method reliably identifies positive cases. Specificity, the mean of 0.788 (SD = 0.137) suggests a moderate to high capability to correctly identify true negatives, though it shows more variability than sensitivity.

PPV and NPV are also noteworthy, with mean values of 0.955 (SD = 0.030) and 0.772 (SD = 0.167), respectively. The PPV indicates a high probability that patients classified as having the syndrome actually do, while the NPV reflects a more variable but substantial probability that patients not classified as having the syndrome truly do not.

Key metrics like the likelihood ratios (LR^+ and LR^-), DOR, and DP further affirm the system's effectiveness. The mean LR^+ was 6.857 (SD = 5.437), indicating a significant increase in the odds of having the syndrome when the test is positive. Conversely, a mean LR^- of 0.077 (SD = 0.086) indicates a low likelihood of having the syndrome when the test is negative. The DOR averaged 171.545 (SD = 136.910), indicating a strong discriminatory test.

Kappa statistic and MCC both averaged 0.728 (SD = 0.182) and 0.730 (SD = 0.181), respectively, suggesting substantial agreement and strong association between predicted and actual classifications. The AUC was 0.867 (SD = 0.087), indicating "very-good" discriminative ability.
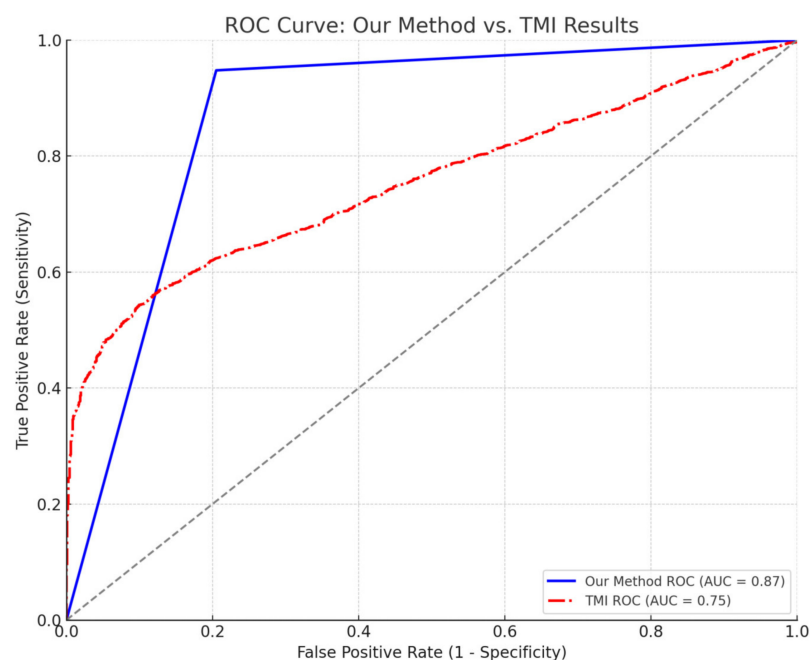
Figure 2. Comparison of ROC curves between our method and the TMI index. The ROC curve for "our method" (blue, solid line) shows an AUC of 0.87. The ROC curve for the TMI index (red, dashed-dot line) is also compared, illustrating its discriminative performance (AUC = 0.75). The diagonal grey dashed line represents the line of no discrimination (AUC = 0.50).

The DP of the classification system averaged 1.955 (SD = 0.619), which falls into the category of "moderate to good" discrimination according to conventional ranges where a DP of 1–2 is considered moderate, and 2–3 is good. This metric underscores the system's ability to effectively differentiate between patients with and without the syndrome.

Overall, the consistency across different test folds is commendable, with key metrics showing relatively low standard deviations, indicating stable performance of the classification system. The results suggest that the method is highly reliable and effective for identifying cardiometabolic syndrome, with strong predictive capabilities and good agreement with actual conditions.

4. Discussion

4.1. Choice of the Classifier

We chose the XGBoost classifier for our binary classification problem due to its ability to handle mixed-type data effectively, which includes features with different measurement scales. XGBoost is an ensemble learning method that uses gradient boosting to optimize prediction performance. It is particularly well suited for problems where feature types and scales vary because it can automatically handle different data types and is less sensitive to feature scaling than other classifiers. Additionally, XGBoost offers strong regularization options to prevent overfitting, and its ability to handle missing data internally further enhances its suitability for our dataset. These advantages, combined with its robust performance in various machine learning tasks, make XGBoost an appropriate choice for our classification problem.

4.2. Comparison with the State-of-the-Art

To our knowledge, this is the first study to classify CMS using a comprehensive set of noninvasive factors in children and adolescents through machine learning techniques. While previous research has explored the relationship between various lifestyle behaviors and cardiometabolic health [52], our study uniquely focuses on CMS—a broader and more encompassing construct than MS—and employs noninvasive measures such as self-rated

health, dietary habits, physical activity, and family history. Additionally, while TMI has been proposed as a similar method for predicting cardiometabolic markers [53], it is a single demographic index. Incorporating a broader range of factors, our method significantly outperformed TMI (adjusted paired *t*-test; $p < 0.05$) (Figure 2; Table 4), providing a more robust and accurate tool for early risk stratification. This novel approach enhances early detection by offering a practical and accessible tool for predicting long-term cardiovascular outcomes in younger populations, thereby addressing a critical gap in the existing literature. Additionally, the large sample size and the relatively homogeneous nature of our population enhanced the statistical significance of the findings.

Few studies have attempted to classify CMS using invasive markers, such as GlycA, L-HDL-PL, and the ApoB/ApoA ratio [17]. While these biomarkers can provide detailed insights into metabolic health, their reliance on blood samples and laboratory analysis makes them less practical for widespread use, especially in large-scale screenings or settings with limited resources. In contrast, our study's noninvasive approach, which utilizes easily accessible indicators like self-rated health and dietary habits, offers a more practical and scalable solution for the early identification of cardiometabolic risks in children and adolescents. Thus, it suits broader public health initiatives and routine clinical assessments.

4.3. Selected Predictors

Research has shown that SRH status significantly influences children and adolescents' CMR. Zhang et al.'s comprehensive meta-analysis demonstrated that higher SRH is associated with more favorable CMR outcomes in this population [54]. Similarly, Näslindh-Ylispangar et al. found a strong correlation between SRH and MS in middle-aged men in Helsinki, highlighting the importance of clinical health status in preventing conditions like diabetes, heart attack, or stroke [55]. Additionally, a cross-sectional study of 533 adolescent girls in the Porto District indicated that better SRH is linked to lower cardiorespiratory risk factors and reduced obesity [56].

Sunlight exposure is a significant factor influencing CMR in children and adolescents. A 2021 study in Iran found that vitamin D deficiency, associated with insufficient sunlight, was linked to higher rates of metabolic syndrome, abdominal obesity, low HDL-C, and high fasting blood sugar in children [57]. Similarly, a 2020 study from Spain demonstrated that vitamin D deficiency correlated with increased BMI, body fat, LDL-C, insulin, and other CMR components in children [58]. Supporting this, a 2016 animal study showed that low-dose ultraviolet radiation (UVR) exposure reduced weight gain and cardiometabolic dysfunction, suggesting that regular, safe sunlight exposure could mitigate obesity risks [59].

Screen time (ST) significantly influences CMR in children and adolescents. Studies have shown that high ST positively correlates with increased obesity, BMI, waist circumference, TG, LDL-C, blood pressure, and lower HDL-C. Research from Iran in 2016 and Brazil in 2017 highlighted the link between high ST and elevated CMR, especially in overweight or obese adolescents [60,61]. Additionally, a 2011 study in the US found that children who watched TV for four or more hours per day had a significantly higher risk of CMR than those who watched less [62].

Consanguineous marriage, or autozygosity, is a significant factor associated with increased CMR among children and adolescents. A 2020 study involving over 10,000 Sub-Saharan African individuals found a strong link between autozygosity and several cardiometabolic traits, particularly in men [63]. Similarly, a 2014 study in Gujarat, India, revealed that children of consanguineous marriages had significantly higher odds of developing heart disease, diabetes, and hypertension [64]. Another 2020 study in Saudi Arabia found that individuals with consanguineous parents were three times more likely to develop obesity compared to those without consanguineous parents [27].

Junk food consumption is strongly associated with increased CMR among children and adolescents, contributing to obesity, hypertension, and diabetes. A 2020 study involving 14,400 students from 30 provinces in Iran found that while junk food was not directly linked to MS incidence, it was associated with higher odds of high blood pressure, elevated

diastolic and systolic blood pressure, overweight, and excess weight [65]. A 2017 meta-analysis of 364 articles confirmed the correlation between unhealthy dietary patterns—characterized by ultra-processed foods, high sodium, fat, and refined carbohydrates—and cardiometabolic alterations in young populations [66]. Additionally, a 2015 study highlighted the link between fast-food consumption and the occurrence of MS and its components in children and adolescents [67].

Our study found that CMR among children and adolescents is gender-dependent. A 2006 meta-analysis highlighted significant differences in MS incidence between males and females, noting greater insulin resistance among girls during puberty and different lipid accumulation patterns between genders [68]. Similarly, a 2018 study observed that boys had higher TG levels, higher SBP, and lower HDL cholesterol compared to girls [69]. Additionally, a 2017 literature review indicated that women are at a higher risk for MS than men [70].

A healthy diet is strongly associated with a reduction in CMR among children and adolescents. A 2022 longitudinal analysis by Peterson et al. found that higher diet quality was linked to improved CMR outcomes among Mexican youth, with at least one CMR factor being positively influenced by better diet quality [71]. Additionally, a meta-analysis confirmed similar findings in adults, demonstrating that healthy dietary patterns—rich in vitamin D, mono- and polyunsaturated fatty acids, fiber, fruits, and vegetables—are associated with better cardiovascular health and lower CMR [72].

Our findings indicate that living in rural or urban areas affects CMR status. Similar studies have shown that the region of residence can significantly impact CMR factors. For example, a 2011 study involving 19,256 Thai adults found a higher prevalence of MS in urban areas compared to rural areas, particularly among men (23.1% vs. 17.9%, $p < 0.05$), with low HDL and hypertriglyceridemia being more common in rural areas [73]. Conversely, a systematic review and meta-analysis of 74,168 American children and adolescents revealed that obesity was more prevalent among those living in rural areas compared to urban counterparts [74]. However, in contrast to these findings, a 2021 study in Eastern Poland found a significant association between MS and rural residency, with rural children exhibiting higher levels of glucose, TGs, blood pressure, and abnormal waist circumference compared to urban residents [75].

Our study showed that the type of dietary fat is strongly linked to CMR in children and adolescents. Supporting this, a systematic review found that replacing saturated fats (SFAs) with monounsaturated (MUFAs) or polyunsaturated fats (PUFAs) lowers total and LDL cholesterol, reducing the risk of CVD, particularly in men [76]. Recent reviews and meta-analyses also confirmed that reducing SFAs and trans-fatty acids (TFAs) intake in youth significantly decreases TG levels and CVD risk later in life [77,78].

Birth order is significantly associated with CMR among children and adolescents. A study involving over 1 million Swedish children found that later-born children had lower BMI than first-borns, though there was no significant link between birth order and BP in young men [79]. Similarly, a 2021 study in Thailand observed that first-borns had a 3.3-times greater risk of obesity compared to later-borns, though no significant associations were found with other cardiometabolic outcomes [80].

Adding extra salt to food is another significant factor linked to CMR in children and adolescents. A review study found a strong positive association between reducing salt and sugar intake and lowering cardiometabolic risks [81]. Similarly, a study in Greece highlighted that reducing added salt intake significantly improves blood pressure control, a key CMR factor in young populations [82].

A family history of NCDs, such as cancer or diabetes, is significantly associated with increased CMR in children and adolescents. A study among Asian Indian adolescents found that those with one or both diabetic parents had higher BMI, waist circumference, impaired glucose levels, hypertension, and lower HDL-C compared to those with parents having normal glucose tolerance [83]. Similarly, a study of African American and Mexican American children aged 12–19 revealed that impaired fasting glucose (IFG), impaired

glucose tolerance (IGT), and prediabetes were prevalent mainly among those with a family history of NCDs, with rates of 25% for type 2 diabetes, 21% for obesity, and 35% for other NCDs [84]. Additionally, a study in Iran involving 14,400 children and adolescents found a significant link between parental NCDs and abnormal CMR in their offspring, particularly in relation to cardiovascular disease, type 2 diabetes, and certain cancers [85].

Tooth health status was the last factor we discussed that significantly impacts CMRs among children and adolescents. Based on a study published in the same region among 5258 students aged 10–18, the authors found an association between tooth brushing frequency and CMRs among adolescents. The results demonstrated that there was an association between higher frequency of tooth brushing and lower mean levels of LDL-C in both genders ($p < 0.0001$), lower frequency of evaluated LDL-C among girls ($p = 0.03$), and lower frequency of elevated BP among boys ($p = 0.03$) [86].

A Japanese 5-year follow-up study found that infrequent tooth brushing was associated with a higher prevalence of MS and CMRs, such as obesity, elevated TGs, high blood pressure, fasting glucose, and reduced HDL cholesterol. Participants who brushed their teeth ≤1 time per day had significantly higher odds of these CMRs compared to those brushing ≥3 times per day (odds ratio = 0.64, 95% confidence interval: 0.45–0.92) [87]. Similarly, a recent study on children aged 12–15 years found that those with poor oral hygiene had higher frequencies of abnormal CMRs, including elevated C-reactive protein, insulin resistance, interleukin-6, and lower adiponectin levels, compared to children with healthier tooth status [88].

4.4. Limitations

A dependable medical diagnosis system must meet specific criteria. It is deemed clinically reliable if it adheres to specific Type I and Type II statistical error rates [89], maintains a false discovery rate (FDR = 1-precision) [90], and achieves DORs [91]. To fulfill these conditions, the system must have minimum sensitivity, specificity, precision, and DOR values of 80%, 95%, 95%, and 100, respectively [92]. Our proposed system meets three conditions: sensitivity, precision, and DOR. However, the system's specificity must be improved to be considered reliable.

All information except laboratory factors used in this study is self-reported by the students and their parents. It might be disposed to recall bias or social desirability bias, especially for the birth-related questions.

One limitation of our study is its cross-sectional design, which can provide valuable insights into associations between various factors and CMR at a specific time. However, it cannot be used to predict future events or establish causality. A cohort study design would be necessary for accurate predictions of future health outcomes, which is the focus of our future work. Despite this limitation, our study offers important and valuable data that contribute to understanding CMR among children and adolescents.

Additionally, due to cross-sectional data collection, we cannot verify the causality between life-related risk factors and CMRs. In addition, we cannot exclude any extra information due to genetic factors. Primary prevention through lifestyle modification should focus on high-risk behavior patterns in children. Given the complex impact of lifestyle factors on CMRs, well-designed, prospective, or genome-wide studies in adolescents are necessary.

5. Conclusions

In conclusion, this study provides a promising advancement in the early identification of CMS in children and adolescents by employing noninvasive factors and machine learning techniques. Addressing the initial research questions, our findings highlight the superior predictive accuracy of our model over traditional methods such as TMI, emphasizing the practical advantages of using easily accessible, noninvasive indicators. This approach not only allows for more effective early intervention strategies but also has broader implications for improving long-term pediatric health outcomes by preventing the progression of CMS

into more severe cardiovascular conditions. Our study underscores the importance of integrating advanced predictive analytics into clinical practice to better manage and reduce the burden of cardiometabolic risks in young populations.

References

1. Zhang, Z.; Fan, S.; Xue, Z.; Yuan, J.; Zhou, Z.; Wang, T.; Liu, J.; Bawudun, A.; Nurmamat, N.; Wang, Y.; et al. Evaluation of the appropriate predictive contributor and diagnostic threshold for the cardio-metabolic syndrome in Chinese Uyghur adults. *BMC Public Health* **2019**, *19*, 613. [CrossRef] [PubMed]
2. Castro, J.P.; El-Atat, F.A.; McFarlane, S.I.; Aneja, A.; Sowers, J.R. Cardiometabolic syndrome: Pathophysiology and treatment. *Curr. Hypertens. Rep.* **2003**, *5*, 393–401. [CrossRef] [PubMed]
3. Srivastava, A.K. Challenges in the treatment of cardiometabolic syndrome. *Indian J. Pharmacol.* **2012**, *44*, 155–156. [CrossRef] [PubMed]
4. Dabelea, D.; Mayer-Davis, E.J.; Saydah, S.; Imperatore, G.; Linder, B.; Divers, J.; Bell, R.; Badaru, A.; Talton, J.W.; Crume, T.; et al. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *Jama* **2014**, *311*, 1778–1786. [CrossRef] [PubMed]
5. Ong, K.L.; Stafford, L.K.; McLaughlin, S.A.; Boyko, E.J.; Vollset, S.E.; Smith, A.E.; Dalton, B.E.; Duprey, J.; Cruz, J.A.; Hagins, H.; et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: A systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **2023**, *402*, 203–234. [CrossRef] [PubMed]
6. Kelli, H.M.; Kassas, I.; Lattouf, O.M. Cardio metabolic syndrome: A global epidemic. *J. Diabetes Metab.* **2015**, *6*, 2. [CrossRef]
7. Sarrafzadegan, N.; Mohammmadifard, N. Cardiovascular Disease in Iran in the Last 40 Years: Prevalence, Mortality, Morbidity, Challenges and Strategies for Cardiovascular Prevention. *Arch. Iran. Med.* **2019**, *22*, 204–210.

8. Shams-Beyranvand, M.; Farzadfar, F.; Naderimagham, S.; Tirani, M.; Maracy, M.R. Estimation of burden of ischemic heart diseases in Isfahan, Iran, 2014: Using incompleteness and misclassification adjustment models. *J. Diabetes Metab. Disord.* **2017**, *16*, 12. [CrossRef]

9. Alberti, K.G.M.M.; Eckel, R.H.; Grundy, S.M.; Zimmet, P.; Fruchart, J.; Smith, S. Harmonizing the Metabolic syndrome. *Circulation* **2009**, *120*, 1640–1645. [CrossRef]

10. Tanrikulu, M.A.; Agirbasli, M.; Berenson, G. Primordial prevention of cardiometabolic risk in childhood. In *Hypertension: From Basic Research to Clinical Practice*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 489–496.

11. Marateb, H.R.; Tasdighi, Z.; Mohebian, M.R.; Naghavi, A.; Hess, M.; Motlagh, M.E.; Heshmat, R.; Mansourian, M.; Mañanas, M.A.; Binder, H.; et al. Classification of psychiatric symptoms using deep interaction networks: The CASPIAN-IV study. *Sci. Rep.* **2021**, *11*, 15706. [CrossRef]

12. Trigka, M.; Dritsas, E. Predicting the Occurrence of Metabolic Syndrome Using Machine Learning Models. *Computation* **2023**, *11*, 170. [CrossRef]

13. Kim, J.; Mun, S.; Lee, S.; Jeong, K.; Baek, Y. Prediction of metabolic and pre-metabolic syndromes using machine learning models with anthropometric, lifestyle, and biochemical factors from a middle-aged population in Korea. *BMC Public Health* **2022**, *22*, 664. [CrossRef] [PubMed]

14. Kupusinac, A.; Doroslovački, R.; Malbaški, D.; Srdić, B.; Stokić, E. A primary estimation of the cardiometabolic risk by using artificial neural networks. *Comput. Biol. Med.* **2013**, *43*, 751–757. [CrossRef]

15. Hollenbeak, C.S.; Spackman, D.E.; Ben-Joseph, R.H.; Okamoto, L.J.; Luce, B.R.; Schwartz, J.S.; Sullivan, S.D. Predicting the prevalence of cardiometabolic risk factors when clinical data are limited. *Value Health* **2007**, *10*, S4–S11. [CrossRef]

16. Ashley-Martin, J.; Ensenauer, R.; Maguire, B.; Kuhle, S. Predicting cardiometabolic markers in children using tri-ponderal mass index: A cross-sectional study. *Arch. Dis. Child.* **2019**, *104*, 577–582. [CrossRef]

17. Ojanen, X.; Cheng, R.; Törmäkangas, T.; Rappaport, N.; Wilmanski, T.; Wu, N.; Fung, E.; Nedelec, R.; Sebert, S.; Vlachopoulos, D.; et al. Towards early risk biomarkers: Serum metabolic signature in childhood predicts cardio-metabolic risk in adulthood. *eBioMedicine* **2021**, *72*, 103611. [CrossRef]

18. Kelishadi, R.; Heidari-Beni, M.; Qorbani, M.; Motamed-Gorji, N.; Motlagh, M.E.; Ziaodini, H.; Taheri, M.; Ahadi, Z.; Aminaee, T.; Heshmat, R. Association between neck and wrist circumferences and cardiometabolic risk in children and adolescents: The CASPIAN-V study. *Nutrition* **2017**, *43–44*, 32–38. [CrossRef]

19. WHO. Physical status: The use and interpretation of anthropometry. In *Report of a WHO Expert Committee*; WHO Technical Report Series 854; World Health Organization: Geneva, Switzerland, 1995.

20. Fredriks, A.M.; van Buuren, S.; Wit, J.M.; Verloove-Vanhorick, S. Body index measurements in 1996–1997 compared with 1980. *Arch. Dis. Child.* **2000**, *82*, 107–112. [CrossRef]

21. Knowles, K.; Paiva, L.; Sanchez, S.; Revilla, L.; Lopez, T.; Yasuda, M.; Yanez, N.; Gelaye, B.; Williams, M. Waist circumference, body mass index, and other measures of adiposity in predicting cardiovascular disease risk factors among Peruvian adults. *Int. J. Hypertens.* **2011**, *2011*, 931402. [CrossRef]

22. Rosner, B.; Cook, N.R.; Daniels, S.; Falkner, B. Childhood blood pressure trends and risk factors for high blood pressure: The NHANES experience 1988–2008. *Hypertension* **2013**, *62*, 247–254. [CrossRef]

23. Zimmet, P.; Alberti, K.G.; Kaufman, F.; Tajima, N.; Silink, M.; Arslanian, S.; Wong, G.; Bennett, P.; Shaw, J.; Caprio, S. The metabolic syndrome in children and adolescents-an IDF consensus report. *Pediatr. Diabetes* **2007**, *8*, 299–306. [CrossRef] [PubMed]

24. Kavey, R.-E.W.; Daniels, S.R.; Lauer, R.M.; Atkins, D.L.; Hayman, L.L.; Taubert, K. American Heart Association guidelines for primary prevention of atherosclerotic cardiovascular disease beginning in childhood. *Circulation* **2003**, *107*, 1562–1566. [CrossRef] [PubMed]

25. Genuth, S.; Alberti, K.; Bennett, P.; Buse, J.; DeFronzo, R.; Kahn, R.; Kitzmiller, J.; Knowler, W.C.; Lebovitz, H.; Lernmark, A. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care* **2003**, *26*, 3160–3168. [PubMed]

26. Soofi, M.; Pasdar, Y.; Karami Matin, B.; Hamzeh, B.; Rezaei, S.; Kazemi Karyani, A.; Moradi Nazar, M.; Soltani, S.; Hajizadeh, M.; Salimi, Y.; et al. Socioeconomic-related inequalities in oral hygiene behaviors: A cross-sectional analysis of the PERSIAN cohort study. *BMC Oral. Health* **2020**, *20*, 63. [CrossRef]

27. Alharbi, K.K.; Al-Sheikh, Y.A.; Alsaadi, M.M.; Mani, B.; Udayaraja, G.K.; Kohailan, M.; Ali Khan, I. Screening for obesity in the offspring of first-cousin consanguineous couples: A Phase-I study in Saudi Arabia. *Saudi J. Biol. Sci.* **2020**, *27*, 242–246. [CrossRef]

28. Meller, F.O.; Loret de Mola, C.; Assunção, M.C.F.; Schäfer, A.A.; Dahly, D.L.; Barros, F.C. Birth order and number of siblings and their association with overweight and obesity: A systematic review and meta-analysis. *Nutr. Rev.* **2018**, *76*, 117–124. [CrossRef]

29. Evensen, E.; Emaus, N.; Kokkvoll, A.; Wilsgaard, T.; Furberg, A.-S.; Skeie, G. The relation between birthweight, childhood body mass index, and overweight and obesity in late adolescence: A longitudinal cohort study from Norway, The Tromsø Study, Fit Futures. *BMJ Open* **2017**, *7*, e015576. [CrossRef]

30. Kelishadi, R.; Farajian, S. The protective effects of breastfeeding on chronic non-communicable diseases in adulthood: A review of evidence. *Adv. Biomed. Res.* **2014**, *3*, 3. [CrossRef]

31. Angoorani, P.; Mostafaei, S.; Kiani, T.; Ejtahed, H.-S.; Motlagh, M.E.; Shafiee, G.; Gorabi, A.M.; Qorbani, M.; Heshmat, R.; Kelishadi, R. Determinants of childhood blood pressure using structure equation model: The CASPIAN–V study. *BMC Cardiovasc. Disord.* **2020**, *20*, 193. [CrossRef]

32. Marateb, H.R.; Mohebian, M.R.; Javanmard, S.H.; Tavallaei, A.A.; Tajadini, M.H.; Heidari-Beni, M.; Mañanas, M.A.; Motlagh, M.E.; Heshmat, R.; Mansourian, M.; et al. Prediction of dyslipidemia using gene mutations, family history of diseases and anthropometric indicators in children and adolescents: The CASPIAN-III study. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 121–130. [CrossRef]

33. Heshmat, R.; Hemati, Z.; Payab, M.; Hamzeh, S.S.; Motlagh, M.E.; Shafiee, G.; Taheri, M.; Ziaodini, H.; Qorbani, M.; Kelishadi, R. Prevalence of different metabolic phenotypes of obesity in Iranian children and adolescents: The CASPIAN V study. *J. Diabetes Metab. Disord.* **2018**, *17*, 211–221. [CrossRef] [PubMed]

34. Khashayar, P.; Kasaeian, A.; Heshmat, R.; Motlagh, M.E.; Mahdavi Gorabi, A.; Noroozi, M.; Qorbani, M.; Kelishadi, R. Childhood overweight and obesity and associated factors in iranian children and adolescents: A multilevel analysis; the CASPIAN-IV study. *Front. Pediatr.* **2018**, *6*, 393. [CrossRef] [PubMed]

35. Shah, B.; Cost, K.T.; Fuller, A.; Birken, C.S.; Anderson, L.N. Sex and gender differences in childhood obesity: Contributing to the research agenda. *BMJ Nutr. Prev. Health* **2020**, *3*, 387. [CrossRef] [PubMed]

36. Angoorani, P.; Mahmoodi, Z.; Ejtahed, H.-S.; Heshmat, R.; Motlagh, M.E.; Qorbani, M.; Kelishadi, R. Determinants of life satisfaction and self-rated health in Iranian children and adolescents: A structure equation model. *BMC Pediatr.* **2022**, *22*, 4. [CrossRef]

37. Pipoyan, D.; Stepanyan, S.; Stepanyan, S.; Beglaryan, M.; Costantini, L.; Molinari, R.; Merendino, N. The Effect of Trans Fatty Acids on Human Health: Regulation and Consumption Patterns. *Foods* **2021**, *10*, 2452. [CrossRef]

38. Unal, I. Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. *Comput. Math. Methods Med.* **2017**, *2017*, 3762651. [CrossRef]

39. Baradaran Mahdavi, S.; Mansourian, M.; Shams, E.; Qorbani, M.; Heshmat, R.; Motlagh, M.E.; Ziaodini, H.; Dashti, R.; Taheri, M.; Kelishadi, R. Association of Sunlight Exposure with Sleep Hours in Iranian Children and Adolescents: The CASPIAN-V Study. *J. Trop. Pediatr.* **2020**, *66*, 4–14. [CrossRef]

40. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006; p. 738.

41. Lafi, S.Q.; Kaneene, J.B. An explanation of the use of principal-components analysis to detect and correct for multicollinearity. *Prev. Vet. Med.* **1992**, *13*, 261–275. [CrossRef]

42. Jakobsen, J.C.; Gluud, C.; Wetterslev, J.; Winkel, P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—A practical guide with flowcharts. *BMC Med. Res. Methodol.* **2017**, *17*, 162. [CrossRef]

43. van Buuren, S.; Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* **2011**, *45*, 1–67. [CrossRef]

44. 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, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

45. Friedman, J.; Hastie, T.; Tibshirani, R. Additive logistic regression: A statistical view of boosting (With discussion and a rejoinder by the authors). *Ann. Stat.* **2000**, *28*, 337–407. [CrossRef]

46. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

47. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for hyper-parameter optimization. *Adv. Neural Inf. Process. Syst.* **2011**, *24*, 2546–2554.

48. Yu, T.; Zhu, H. Hyper-parameter optimization: A review of algorithms and applications. *arXiv* **2020**, arXiv:2003.05689.

49. Mansourian, M.; Marateb, H.R.; Mansourian, M.; Mohebbian, M.R.; Binder, H.; Mañanas, M.Á. Rigorous performance assessment of computer-aided medical diagnosis and prognosis systems: A biostatistical perspective on data mining. In *Modelling and Analysis of Active Biopotential Signals in Healthcare*; IOP Publishing: Bristol, UK, 2020; Volume 2. [CrossRef]

50. *ISO/IEC TS 4213*; Information Technology—Artificial Intelligence—Assessment of Machine Learning Classification Performance. International Organization for Standardization: Geneva, Switzerland, 2022.

51. Kuncheva, L. Fundamentals of Pattern Recognition. In *Combining Pattern Classifiers*; Kuncheva, L., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2004; pp. 1–43.

52. Forseth, B.; Noel-MacDonnell, J.R.; Hampl, S.; Carlson, J.A.; Halpin, K.; Davis, A.; Phillips, T.; Shook, R.P. Relationship between youth cardiometabolic health and physical activity in medical records. *PLoS ONE* **2024**, *19*, e0303583. [CrossRef] [PubMed]

53. Gul Siraz, U.; Hatipoglu, N.; Mazicioglu, M.M.; Ozturk, A.; Cicek, B.; Kurtoglu, S. Triponderal mass index is as strong as body mass index in the determination of obesity and adiposity. *Nutrition* **2023**, *105*, 111846. [CrossRef]

54. Zhang, T.; Lu, G.; Wu, X.Y. Associations between physical activity, sedentary behaviour and self-rated health among the general population of children and adolescents: A systematic review and meta-analysis. *BMC Public Health* **2020**, *20*, 1343. [CrossRef]

55. Näslindh-Ylispangar, A.; Sihvonen, M.; Vanhanen, H.; Kekki, P. Self-rated health and risk factors for metabolic syndrome among middle-aged men. *Public Health Nurs.* **2005**, *22*, 515–522. [CrossRef]

56. Mota, J.; Santos, R.M.; Silva, P.; Aires, L.; Martins, C.; Vale, S. Associations between self-rated health with cardiorespiratory fitness and obesity status among adolescent girls. *J. Phys. Act. Health* **2012**, *9*, 378–381. [CrossRef]

57. Qorbani, M.; Heidari-Beni, M.; Ejtahed, H.-S.; Shafiee, G.; Goodarzi, F.; Tamehri Zadeh, S.S.; Khademian, M.; Mohammadian Khonsari, N.; Motlagh, M.E.; Asayesh, H. Association of vitamin D status and cardio-metabolic risk factors in children and adolescents: The CASPIAN-V study. *BMC Nutr.* **2021**, *7*, 71. [CrossRef]

58. Durá-Travé, T.; Gallinas-Victoriano, F.; Peñafiel-Freire, D.M.; Urretavizcaya-Martinez, M.; Moreno-González, P.; Chueca-Guindulain, M.J. Hypovitaminosis D and cardiometabolic risk factors in adolescents with severe obesity. *Children* **2020**, *7*, 10. [CrossRef] [PubMed]

59. Fleury, N.; Geldenhuys, S.; Gorman, S. Sun exposure and its effects on human health: Mechanisms through which sun exposure could reduce the risk of developing obesity and cardiometabolic dysfunction. *Int. J. Environ. Res. Public Health* **2016**, *13*, 999. [CrossRef] [PubMed]

60. Heshmat, R.; Qorbani, M.; Shahr Babaki, A.E.; Djalalinia, S.; Ataei-Jafari, A.; Motlagh, M.E.; Ardalan, G.; Arefirad, T.; Rezaei, F.; Asayesh, H. Joint association of screen time and physical activity with cardiometabolic risk factors in a national sample of Iranian adolescents: The CASPIANIII study. *PLoS ONE* **2016**, *11*, e0154502. [CrossRef]

61. Cureau, F.; Ekelund, U.; Bloch, K.; Schaan, B. Does body mass index modify the association between physical activity and screen time with cardiometabolic risk factors in adolescents? Findings from a country-wide survey. *Int. J. Obes.* **2017**, *41*, 551–559. [CrossRef] [PubMed]

62. Carson, V.; Janssen, I. Volume, patterns, and types of sedentary behavior and cardio-metabolic health in children and adolescents: A cross-sectional study. *BMC Public Health* **2011**, *11*, 274. [CrossRef]

63. Ceballos, F.C.; Hazelhurst, S.; Clark, D.W.; Agongo, G.; Asiki, G.; Boua, P.R.; Xavier Gómez-Olivé, F.; Mashinya, F.; Norris, S.; Wilson, J.F. Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. *Nat. Commun.* **2020**, *11*, 5754. [CrossRef]

64. Bhasin, P.; Kapoor, S. Impact of consanguinity on cardio-metabolic health and other diseases: Findings from an Afro-Indian tribal community. *J. Community Genet.* **2015**, *6*, 129–135. [CrossRef]

65. Azemati, B.; Kelishadi, R.; Ahadi, Z.; Shafiee, G.; Taheri, M.; Ziaodini, H.; Qorbani, M.; Heshmat, R. Association between junk food consumption and cardiometabolic risk factors in a national sample of Iranian children and adolescents population: The CASPIAN-V study. *Eat. Weight. Disord. EWD* **2020**, *25*, 329–335. [CrossRef]

66. Rocha, N.P.; Milagres, L.C.; Longo, G.Z.; Ribeiro, A.Q.; Novaes, J.F. Association between dietary pattern and cardiometabolic risk in children and adolescents: A systematic review. *J. Pediatr.* **2017**, *93*, 214–222. [CrossRef]

67. Asghari, G.; Yuzbashian, E.; Mirmiran, P.; Mahmoodi, B.; Azizi, F. Fast Food Intake Increases the Incidence of Metabolic Syndrome in Children and Adolescents: Tehran Lipid and Glucose Study. *PLoS ONE* **2015**, *10*, e0139641. [CrossRef]

68. Regitz-Zagrosek, V.; Lehmkuhl, E.; Mahmoodzadeh, S. Gender Aspects of the Role of the Metabolic Syndrome as a Risk Factor for Cardiovascular Disease. *Gend. Med.* **2007**, *4* (Suppl. B), S162–S177. [CrossRef] [PubMed]

69. Barstad, L.H.; Júlíusson, P.B.; Johnson, L.K.; Hertel, J.K.; Lekhal, S.; Hjelmesæth, J. Gender-related differences in cardiometabolic risk factors and lifestyle behaviors in treatment-seeking adolescents with severe obesity. *BMC Pediatr.* **2018**, *18*, 61. [CrossRef] [PubMed]

70. Pucci, G.; Alcidi, R.; Tap, L.; Battista, F.; Mattace-Raso, F.; Schillaci, G. Sex- and gender-related prevalence, cardiovascular risk and therapeutic approach in metabolic syndrome: A review of the literature. *Pharmacol. Res.* **2017**, *120*, 34–42. [CrossRef] [PubMed]

71. Aljahdali, A.A.; Peterson, K.E.; Cantoral, A.; Ruiz-Narvaez, E.; Tellez-Rojo, M.M.; Kim, H.M.; Hébert, J.R.; Wirth, M.D.; Torres-Olascoaga, L.A.; Shivappa, N.; et al. Diet Quality Scores and Cardiometabolic Risk Factors in Mexican Children and Adolescents: A Longitudinal Analysis. *Nutrients* **2022**, *14*, 896. [CrossRef]

72. Funtikova, A.N.; Navarro, E.; Bawaked, R.A.; Fíto, M.; Schröder, H. Impact of diet on cardiometabolic health in children and adolescents. *Nutr. J.* **2015**, *14*, 118. [CrossRef]

73. Aekplakorn, W.; Kessomboon, P.; Sangthong, R.; Chariyalertsak, S.; Putwatana, P.; Inthawong, R.; Nitiyanant, W.; Taneepanichskul, S.; The NHES IV Study Group. Urban and rural variation in clustering of metabolic syndrome components in the Thai population: Results from the fourth National Health Examination Survey 2009. *BMC Public Health* **2011**, *11*, 854. [CrossRef]

74. Johnson III, J.A.; Johnson, A.M. Urban-rural differences in childhood and adolescent obesity in the United States: A systematic review and meta-analysis. *Child. Obes.* **2015**, *11*, 233–241. [CrossRef]

75. Nowicki, G.J.; Ślusarska, B.; Naylor, K.; Prystupa, A.; Rudnicka-Drożak, E.; Halyuk, U.; Pokotylo, P. The relationship between the metabolic syndrome and the place of residence in the local community on the example of the janów lubelski district in eastern Poland: A population-based study. *Diabetes Metab. Syndr. Obes.* **2021**, *14*, 2041–2056. [CrossRef]

76. Schwab, U.; Lauritzen, L.; Tholstrup, T.; Haldorsson, T.I.; Riserus, U.; Uusitupa, M.; Becker, W. Effect of the amount and type of dietary fat on cardiometabolic risk factors and risk of developing type 2 diabetes, cardiovascular diseases, and cancer: A systematic review. *Food Nutr. Res.* **2014**, *58*, 25145. [CrossRef]

77. Morenga, L.T.; Montez, J.M. Health effects of saturated and trans-fatty acid intake in children and adolescents: Systematic review and meta-analysis. *PLoS ONE* **2017**, *12*, e0186672. [CrossRef]

78. Micha, R.; Mozaffarian, D. Saturated fat and cardiometabolic risk factors, coronary heart disease, stroke, and diabetes: A fresh look at the evidence. *Lipids* **2010**, *45*, 893–905. [CrossRef] [PubMed]

79. Jelenkovic, A.; Silventoinen, K.; Tynelius, P.; Myrskylä, M.; Rasmussen, F. Association of birth order with cardiovascular disease risk factors in young adulthood: A study of one million Swedish men. *PLoS ONE* **2013**, *8*, e63361. [CrossRef] [PubMed]

80. Aurpibul, L.; Butler, É.M.; Wongthanee, A.; Rerkasem, A.; Pruenglampoo, S.; Mangklabruks, A.; Rerkasem, K.; Derraik, J.G. Birth order is associated with an increased risk of obesity in young adults in Thailand. *J. Epidemiol. Community Health* **2021**, *75*, 305–308. [CrossRef] [PubMed]

81. Ponzo, V.; Pellegrini, M.; Costelli, P.; Vázquez-Araújo, L.; Gayoso, L.; D'Eusebio, C.; Ghigo, E.; Bo, S. Strategies for reducing salt and sugar intakes in individuals at increased cardiometabolic risk. *Nutrients* **2021**, *13*, 279. [CrossRef] [PubMed]

82. Tragomalou, A.; Moschonis, G.; Manios, Y.; Kassari, P.; Ioakimidis, I.; Diou, C.; Stefanopoulos, L.; Lekka, E.; Maglaveras, N.; Delopoulos, A. Novel e-health applications for the management of cardiometabolic risk factors in children and adolescents in Greece. *Nutrients* **2020**, *12*, 1380. [CrossRef]

83. Anjana, R.M.; Lakshminarayanan, S.; Deepa, M.; Farooq, S.; Pradeepa, R.; Mohan, V. Parental history of type 2 diabetes mellitus, metabolic syndrome, and cardiometabolic risk factors in Asian Indian adolescents. *Metabolism* **2009**, *58*, 344–350. [CrossRef]

84. Li, C.; Ford, E.S.; Zhao, G.; Mokdad, A.H. Prevalence of pre-diabetes and its association with clustering of cardiometabolic risk factors and hyperinsulinemia among U.S. adolescents: National Health and Nutrition Examination Survey 2005–2006. *Diabetes Care* **2009**, *32*, 342–347. [CrossRef]

85. Ejtahed, H.-S.; Heshmat, R.; Motlagh, M.E.; Hasani-Ranjbar, S.; Ziaodini, H.; Taheri, M.; Ahadi, Z.; Aminaee, T.; Shafiee, G.; Goodarzi, A. Association of parental obesity with cardiometabolic risk factors in their children: The CASPIAN-V study. *PLoS ONE* **2018**, *13*, e0193978. [CrossRef]

86. Kelishadi, R.; Mirmoghtadaee, P.; Qorbani, M.; Motlagh, M.E.; Heshmat, R.; Taslimi, M.; Mahmoudarabi, M.; Ardalan, G.; Larijani, B. Tooth brushing and cardiometabolic risk factors in adolescents: Is there an association? The CASPIAN-III study. *Int. J. Prev. Med.* **2013**, *4*, 271.

87. Tanaka, A.; Takeuchi, K.; Furuta, M.; Takeshita, T.; Suma, S.; Shinagawa, T.; Shimazaki, Y.; Yamashita, Y. Relationship of toothbrushing to metabolic syndrome in middle-aged adults. *J. Clin. Periodontol.* **2018**, *45*, 538–547. [CrossRef]

88. Baskaradoss, J.K.; Tavares, M.; Al-Mulla, F.; Al-Ozairi, E.; Abu-Farha, M.; Bin-Hasan, S.; Alsumait, A.; Devarajan, S.; Alqaderi, H. Association between Frequency of Toothbrushing and Metabolic Syndrome among Adolescents: A 5-Year Follow-Up Study. *Int. J. Environ. Res. Public Health* **2022**, *19*, 508. [CrossRef]

89. Ellis, P.D. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*; Cambridge University Press: Cambridge, UK, 2010.

90. Colquhoun, D. An investigation of the false discovery rate and the misinterpretation of *p*-values. *R. Soc. Open Sci.* **2014**, *1*, 140216. [CrossRef]

91. Guilherme, H.O.; Nesbitt, G.C.; Murphy, J.G.; Habermann, T.M. *Mayo Clinic Medical Manual and Mayo Clinic Internal Medicine Review*; CRC Press: Boca Raton, FL, USA, 2007.

92. Ghosh, A.K. *Mayo Clinic Internal Medicine Review*, 8th ed.; mcGraw-Hill: New York, NY, USA, 2008; 1101p, ISBN 978-1420084788.