ISRG Journal of Multidisciplinary Studies (ISRGJMS)

ISRG PUBLISHERS

Abbreviated Key Title: isrg j. multidiscip. Stud.

ISSN: 2584-0452 (Online)

Journal homepage: https://isrgpublishers.com/isrgjms/

Volume – IV, Issue – VIII (August) 2026

Frequency: Monthly

Adolescent Metabolic Syndrome with Age- and Sex-Specific Thresholds: A Laboratory-Enhanced, Temporally Validated Machine Learning Model

Saqlain Abbas[1], Sibtain Abbas[1*], Imtiaz Hussain[1]

[1] School of Mathematics, Hohai University, Nanjing, 210098, China.

| **Received:** 10.08.2026 | **Accepted:** 16.08.2026 | **Published:** 18.08.2026

***Corresponding author:** Sibtain Abbas

Abstract

Background: *Adolescent metabolic syndrome (MetS) represents an important early-life cardiometabolic risk factor for type 2 diabetes and cardiovascular disease, yet its identification remains inconsistent because no internationally accepted paediatric diagnostic standard exists. Many studies default to fixed adult anthropometric thresholds, such as waist circumference cut-offs of ≥102 cm (men) and ≥88 cm (women), despite body composition changing substantially and non-linearly across puberty, systematically altering who is classified as MetS-positive in youth.*

Objective: *To define adolescent MetS using age- and sex-specific percentile thresholds (Johnson/Ford modified paediatric ATP III criteria, with blood pressure further stratified by age, sex, and height) rather than fixed adult cut-offs, and, on this paediatric-appropriate foundation, to develop and temporally validate a laboratory-enhanced machine learning model evaluated on a post-pandemic holdout absent from prior work with formal ablation to quantify the contribution of each predictor domain.*

Methods: *Seven pooled NHANES survey cycles (2007–2023) yielded 3,567 adolescents aged 12–19 years with complete diagnostic data, the longest observation window applied to this question to date. MetS was classified using age- and sex-specific 90th percentile thresholds for waist circumference and age-, sex-, and height-specific 90th percentile thresholds for blood pressure. A laboratory-enriched predictor panel spanning insulin resistance, lipid metabolism, purine metabolism, and adiposity was evaluated through a two-stage feature selection process (Elastic Net followed by RFECV) and a four-strategy imbalance-handling ablation, yielding a final eight-variable LightGBM classifier with SMOTE-based correction. The model was trained exclusively on 2007-2020 data (n=3,150), with the August 2021-August 2023 cycle (n = 417) held out entirely as a temporally independent test set following the COVID-19-related interruption in NHANES data collection.*

DOI: 10.5281/zenodo.21992844

Results: *The age- and sex-specific percentile definition identified 20.6% more MetS-positive adolescents than an adult-derived waist-circumference definition applied to the same sample (7.88% vs 6.25% prevalence), confirming that percentile-based paediatric criteria capture a materially larger and more physiologically appropriate at-risk population. On the temporal holdout, the resulting model achieved excellent discrimination (ROC-AUC 0.977, 95% CI 0.959–0.991; PR-AUC 0.868, 0.774–0.943) and a negative predictive value of 0.979, correctly excluding MetS in approximately 98% of screened adolescents. SHAP interpretation and systematic ablation converged on a consistent biological signature: TG/HDL-C ratio, BMI, and HOMA-IR jointly drove predictive performance, with the lipid domain alone accounting for the majority of discriminative signal (PR-AUC fell by 0.373 upon removal), while removing the insulin-resistance domain produced only marginal loss, indicating overlapping rather than absent biological signal. Decision curve analysis confirmed net clinical benefit across screening-relevant threshold probabilities (5–50%).*

Conclusion: *Defining adolescent MetS using age- and sex-specific percentile thresholds, rather than adult cut-offs, relevantly changes who is detected as at risk and provides a more physiologically valid foundation for risk prediction. Built on this definition, a laboratory-enhanced machine learning model validated on a temporally independent August 2021-August 2023 cohort with formal ablation of each predictor domain sustained high, clinically meaningful discriminative accuracy, offering a more rigorous alternative to prior adult-threshold, random-split adolescent MetS models and supporting its potential as a population-level screening tool pending external, prospective validation.*

Keywords: *Age- and sex-specific percentile criteria; Adolescent metabolic syndrome; Paediatric diagnostic thresholds; Machine learning; LightGBM; NHANES; Temporal validation; SHAP*

1. Introduction

Adolescent metabolic syndrome (MetS) is a clustering of cardiometabolic abnormalities including adiposity, dyslipidaemia, elevated blood pressure and dysglycaemia. It is a well-established early-life precursor to type 2 diabetes that is no longer confined to adulthood [1, 2]. Rising rates of paediatric obesity have made adolescent MetS an increasingly common and clinically important finding; however, its identification in youth remains inconsistent across studies because, unlike in adult MetS, no single internationally accepted diagnostic definition exists for adolescents [3]. A crucial methodological problem is that many studies applying MetS criteria to adolescents default to fixed adult thresholds like waist circumference cut-offs of ≥102 cm (men) and ≥88 cm (women) [4, 5], even though body composition changes substantially and non-linearly across puberty. Applying adult anthropometric thresholds to a growing body systematically modifies who is classified as MetS-positive. In a nationally representative NHANES sample, adult-derived waist cut-offs would classify roughly one-fifth fewer adolescents as MetS-positive than age- and sex-specific paediatric percentile criteria, suggesting that adult thresholds meaningfully minimize adolescent MetS burden. Paediatric-specific criteria, such as the Johnson/Ford modification of the ATP III definition, are defined using age- and sex-specific percentile thresholds for waist circumference and for blood pressure using age-, sex-, and height-specific percentile thresholds, offering a more physiologically appropriate diagnostic standard, though they remain underutilized in the predictive modelling literature [6, 7].

Simultaneously, machine learning techniques have increasingly been applied to MetS risk prediction, inspired by the syndrome's multifactorial, non-linear presentation. Ensemble tree-based methods such as gradient boosting have shown strong discriminative performance in several NHANES-based analyses. However, three limitations recur across this literature. Firstly, many models are validated using random train-test splits within a single pooled dataset, which evaluates a model's ability to interpolate within a fixed time window rather than its ability to generalize to a certain future population sampled under different survey and public health conditions. This distinction became especially relevant following the effects of the COVID-19 pandemic on adolescent health behaviours and healthcare access [8]. Secondly, several existing models, including a recent LightGBM-based adolescent MetS model trained on five NHANES survey periods (2007–2016) with 139 MetS-positive cases, rely on adult anthropometric MetS criteria rather than paediatric-specific thresholds, and restrict predictors to anthropometric and demographic variables, excluding laboratory-derived biomarkers of insulin resistance and lipid metabolism that are structurally central to MetS pathophysiology [9]. Lastly, few studies pair predictive performance with formal ablation analysis to quantify how much discriminative signal is actually available to each predictor domain, leaving the biological interpretability of "black-box" model output largely unexamined.

The present study addresses these gaps using seven NHANES survey periods spanning 2007–2023, the longest observation window applied to this question to date, in a nationally representative sample of adolescents aged 12-19 years. MetS was defined using paediatric-appropriate Johnson/Ford criteria rather than adult thresholds. A laboratory-enhanced predictor panel spanning insulin resistance (HOMA-IR, log-insulin), lipid metabolism (TG/HDL-C ratio, total cholesterol), purine metabolism (uric acid), and adiposity (body mass index [BMI]) was evaluated using a light gradient boosting machine (LightGBM) classifier with SMOTE-based imbalance correction, selected through a systematic two-stage feature selection process and a four-strategy imbalance-handling ablation. Critically, model development, feature selection, imputation, and threshold optimization were performed exclusively on NHANES survey periods 2007–2020, with the August 2021-August 2023 period held out entirely as a temporally independent test set following the COVID-19-related interruption in NHANES data collection. This design provides a more stringent and clinically meaningful test of prospective applicability than a randomly split internal validation.

DOI: 10.5281/zenodo.21992844

At the same time, SHAP-based interpretation and systematic ablation experiments quantify the contribution of individual predictors and modelling choices to overall performance. This clarifies, for instance, the extent to which the lipid domain versus the insulin resistance domain drives predictive accuracy.

2. Related Work

2.1 Evolution and heterogeneity of paediatric MetS definitions

None of the paediatric MetS definitions has achieved consensus yet [3]. Unlike the adult ATP III/AHA-NHLBI [4, 5, 10] and adult IDF [11] definitions, Cook et al. first adapted NCEP/ATP III criteria for adolescents [12], and de Ferranti et al. later modified a similar definition but with lower waist-circumference and lipid thresholds [13], and Weiss et al. used a BMI-based approach [14], which tends to generate higher prevalence estimates. The IDF subsequently introduced an adult-derived definition stratified by age band [15, 16]. This definitional heterogeneity has direct empirical consequences: comparative studies applying multiple definitions to the same cohort have reported paediatric MetS prevalence ranging from roughly 2% to over 40% depending on which criteria were used [17]. Early NHANES-based work using age-modified ATP III criteria found adolescent MetS prevalence rising from 4.2% in NHANES III to 6.4% in NHANES 1999–2000 [18], emphasizing both temporal drift and definitional sensitivity in any reported estimate. This body of work motivates the present study's use of an explicit, percentile-based paediatric definition Johnson/Ford [6, 7] rather than adult thresholds, and the direct empirical quantification of how an adult-cut-off definition would shift case counts within the same analytic sample.

2.2 Machine learning approaches to MetS prediction.

Tree-ensemble approaches have become an effective approach to MetS-related risk prediction in NHANES data given the syndrome's non-linear, multi-component arrangement [9, 19]. Gradient boosted models paired with SHAP interpretation have reported strong discrimination for MetS and closely related cardiometabolic outcomes in adult NHANES cohorts, typically in the ROC-AUC range of 0.87-0.94. Within the paediatric NHANES literature specifically, the most directly comparable prior work is a LightGBM-based adolescent MetS model built on NHANES 2007–2016 [9], which is compared in detail with the present results. Briefly, that model relied on a randomly split rather than temporally held-out test set, applied an adult-derived rather than paediatric waist criterion, and restricted its predictor set to anthropometric and demographic variables, excluding laboratory-derived insulin-resistance and lipid indices that are included in the present study [9].

2.3 Derived metabolic indices in ML pipelines

It has been suggested that surrogate indices such as the TyG index and TG/HDL-C ratio have gained traction as low-cost measures for insulin resistance and atherogenic dyslipidaemia [20, 21] in ML pipelines for related cardiometabolic outcomes. A recent ensemble learning comparison for adolescent NAFLD risk benchmarked TyG-based logistic regression against tree ensembles and found that SHAP identified waist circumference, triglycerides, insulin, and HDL as leading predictors with non-linear threshold effects [22], a pattern consistent with the dyslipidaemia insulin resistance focus identified by SHAP in the present study and one that motivated including both TyG index and ALT/GGT ratio as candidate predictors here, even though neither survived feature selection in the final model.

2.4 Synthesis and remaining gaps

This literature establishes that (i) definitional choice materially alters who is classified as MetS-positive in paediatric populations, (ii) gradient-boosted ensembles paired with SHAP now represent the standard approach for MetS-related prediction in NHANES data, and (iii) the most comparable existing adolescent MetS model achieves powerful discrimination but does so under a random-split design, an adult-derived waist criterion, and an anthropometry-only predictor set. No prior study, to our knowledge, has combined a paediatric-appropriate MetS definition, a laboratory-enhanced predictor panel spanning insulin resistance and lipid metabolism, a strictly temporal (pre- versus post-pandemic) validation split, and formal ablation experiments quantifying the incremental value of each of the predictor domains, which is the contribution of the present work.

3. Method

3.1 Data source and study population

The data used in this study were obtained from the National Health and Nutrition Examination Survey (NHANES), a nationally representative source provided by the National Center for Health Statistics (NCHS) of the U.S. Centers for Disease Control and Prevention (CDC). NHANES uses a complex multi-stage probability sampling design that combines household interviews, laboratory tests, standardized physical examinations, and dietary questionnaires [23]. NCHS collects, shares, and analyses relevant and reliable health information to understand health problems, keep the public informed, and support decisions that improve the nation's health. Data from seven NHANES survey periods were pooled for the analysis: 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, 2017-March 2020, and 2021-2023 [23, 24]. NHANES data are publicly available data files and do not include identifying information, so this secondary analysis did not need additional institutional review board approval. Participants were eligible if they were aged 12–19 years, were included in the fasting laboratory subsample, and had sufficient information to classify all five MetS diagnostic components. (waist Circumference, blood pressure, triglycerides, HDL-C, and fasting plasma glucose). The study population was restricted to adolescents aged 12–19 years; 10,178 adolescents were identified across the seven NHANES survey periods, and 3,567 had complete diagnostic data and were included in the final analytic sample. Participants lacking data on any of the five diagnostic component variables required to classify MetS were excluded from the analytic sample before modelling.

3.2 Variable harmonization across NHANES survey periods

Because NHANES variable names and measurement protocols differed across survey cycles, all variables were harmonized before pooling. Blood pressure variable names were mapped from cycle-specific names (BPXSY1/BPXOSY1, BPXDI1/BPXODI1) into common systolic and diastolic blood pressure columns. Fasting subsample weights were similarly unified by combining WTSAF2YR from the 2007–2016 cycles and WTSAFPRP from the 2017–2023 periods into one harmonised weight variable. Race and ethnicity were harmonised from the five-category coding used in earlier cycles to the six-category coding introduced in 2011–2012 using a validated mapping. C-reactive protein (CRP) was

DOI: 10.5281/zenodo.21992844

sourced from LBXCRP reported in (mg/dL) in the 2007–2010 cycles and converted to (mg/L) by multiplying by 10; later cycles used LBXHSCRP, which is already in (mg/L). CRP was harmonised where available but was not included in the primary modelling analysis because it was unavailable in the 2011-2014 survey periods and had substantial cycle-specific missingness.

3.3 Derived biomarkers

Six laboratory-derived biomarkers were calculated from the original NHANES laboratory measurements. HOMA-IR was calculated using the standard formula for U.S. laboratory units [25] as:

$$HOMA - IR = \frac{fasting\ insulin\ (\mu IU/mL) \times fasting\ glucose\ (mg/dL)}{405}$$

Log-insulin was calculated as the natural logarithm of fasting insulin to reduce the influence of its heavily right-skewed distribution:

$$Log\ insulin = \ln(\text{Fasting insulin } (\mu IU/mL))$$

The TG/HDL-C ratio was calculated as serum triglycerides divided by HDL-cholesterol [21]:

$$TG/HDL - C\ ratio = \frac{triglycerides\ (mg/dL)}{HDL - C\ (mg/dL)}$$

The triglyceride-glucose index was calculated as [20]:

$$TyG\ index = ln\ (\frac{triglyceride(mg/dL) \times fasting\ glucose\ (mg/dL)}{2})$$

The ALT/GGT ratio was calculated as serum alanine aminotransferase divided by gamma-glutamyl transferase:

$$ALT/GGT\ ratio = \frac{ALT\ (U/L)}{GGT\ (U/L)}$$

Vitamin D was converted from nmol/L to ng/mL as:

$$Vitamin\ D\ (ng/mL) = \frac{Vitamin\ D\ \ (nmol/L)}{2.496}$$

3.4 Metabolic Syndrome outcome definition

Metabolic Syndrome was defined using the Johnson/Ford-modified Paediatric ATP III criteria [6, 7], which apply age- and sex-specific 90th-percentile thresholds, while elevated blood pressure was defined using age-, sex-, and height-specific 90th-percentile thresholds. A participant was classified as MetS-positive if three or more of the following five components were present:

- Abdominal obesity was defined as waist circumference ≥ the age- and sex-specific 90th percentile;
- Systolic or diastolic blood pressure ≥ age-, sex-, and height-specific 90th percentile, using NHLBI/NHBPEP paediatric reference table [26]. For participants who are aged 18-19 years, the age 17 reference row was applied, as published paediatric tables do not extend beyond 17;
- Serum triglycerides ≥ 110 mg/dL;
- HDL-cholesterol ≤ 40 mg/dL;
- Fasting plasma glucose ≥ 100 mg/dL.

This MetS definition used paediatric waist circumference criteria rather than fixed adult AHA/NHLBI cut-offs. Abdominal obesity was therefore classified using age- and sex-specific 90th-percentile waist circumference values, rather than adult thresholds of ≥102 cm for males and ≥88 cm for females. To verify the diagnostic rule, a reproducibility audit was performed; the recalculated component totals matched the final MetS classification for all 3,567 participants, with no mismatches observed.

3.5 Temporal Validation design

A temporal validation design was used to assess model generalisation to a later NHANES survey period. NHANES survey periods from 2007-2008 through 2017–March 2020 formed the development set (n = 3,150; MetS-positive n = 244; prevalence 7.75%). The 2021–2023 survey period was held out as the temporal test set (n = 417; MetS-positive n = 37; prevalence 8.87%) and was not used during model training, feature selection, imputation fitting, hyperparameter selection, or threshold optimisation. This design evaluated whether the model maintained its performance during a temporally independent August 2021-August 2023 survey period, collected after NHANES field operations resumed following the COVID-19-related suspension [24]. Baseline characteristics of the development and temporal test sets were compared descriptively and are reported in **Table 2**.

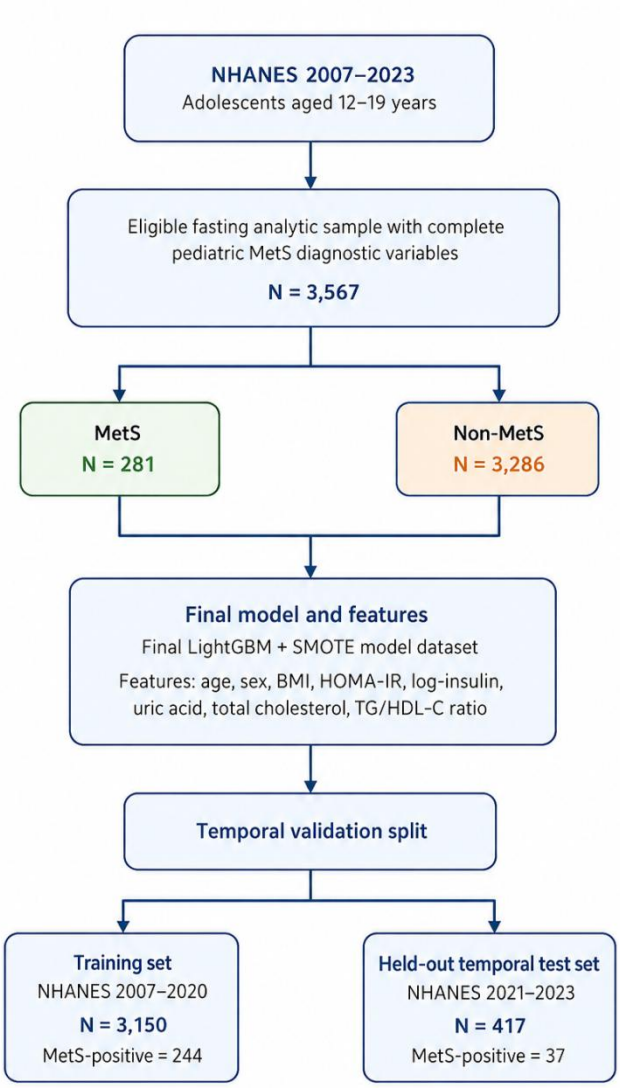

Figure 1. Study cohort selection, metabolic syndrome classification, feature set, and temporal validation design. Adolescents aged 12–19 years from NHANES 2007–2023 were screened to identify an eligible fasting analytic sample with complete paediatric metabolic syndrome (MetS) diagnostic variables (N = 3,567). Participants were classified as MetS (n =

DOI: 10.5281/zenodo.21992844

281) or non-MetS (n = 3,286). The final LightGBM model with SMOTE incorporated age, sex, BMI, HOMA-IR, log-transformed insulin, uric acid, total cholesterol, and TG/HDL-C ratio. Temporal validation was performed by dividing the cohort into a training set from NHANES 2007-2020 (n = 3,150; 244 MetS-positive) and a held-out temporal test set from NHANES 2021-2023 (n = 417; 37 MetS-positive).

3.6 Missing data handling

Participants with missing data for any diagnostic variable needed to classify MetS status were excluded from the analytic sample before modelling. For continuous predictor variables, missing values were addressed using iterative multivariate imputation implemented with scikit-learn's IterativeImputer [27] (max iteration = 10, random_state = 42, initial strategy = median). The imputation step was placed at the beginning of each modelling pipeline, fitted only on the development set, and then applied to the temporal test set through the fitted pipeline to prevent information leakage. Vitamin D was structurally unavailable in the 2017–March 2020 NHANES period and was therefore evaluated separately in a pre-specified sensitivity analysis among participants with available vitamin D measurements.

3.7 Feature Selection

TyG index and ALT/GGT ratio were calculated as prespecified candidate-derived biomarkers and were evaluated in the feature selection audit, but neither was included in the final predictor set after the Elastic Net/RFECV selection and model-locking process. Feature selection was conducted in a two-stage approach on the development set only. In the first stage, Elastic Net [28] regularised logistic regression was used with 5-fold cross-validation on a set of 19 candidate predictors representing laboratory, anthropometric, and demographic domains. Predictors that maintained non-zero coefficients across the regularization process were carried forward. The second stage employed recursive feature elimination with cross-validation [27, 29] (RFECV) with the candidate set, where feature importance is calculated by PR-AUC on the inner CV folds and the LightGBM estimator is used as the explainer. A bootstrap stability audit was also conducted using 100 resampled training sets to verify the stability of the selection. All continuous predictors were centered and standardized before evaluating multicollinearity through the variance inflation factor (VIF). There was no significant multicollinearity, as all eight variables in the final model had VIF values below 3.0, ranging from 1.065 to 2.985. The final predictor set was locked, including age, sex, BMI, HOMA-IR, log-insulin, uric acid, total cholesterol, and TG/HDL-C ratio.

3.8 Model development and class imbalance handling

A four-strategy ablation study was conducted to evaluate class imbalance approaches:

- no adjustment
- class weighting only (balanced)
- SMOTE only
- SMOTE + class weighting

SMOTE-only [30] using a minority to majority sampling ratio = 0.50 achieved the best combination of PR-AUC and the lowest Brier score simultaneously and was selected as the primary strategy. SMOTE was applied only during the model training pipeline and was not used on the temporal test set. The final classifier was a Light Gradient Boosting Machine (LightGBM) [19] trained with fixed hyperparameters: n_estimators = 300, learning_rate = 0.03, num_leaves = 15, min_child_samples = 20, subsample = 0.90, reg_lambda = 1.0, and random_state = 42. The complete modelling pipeline was: IterativeImputer → SMOTE → LightGBM. A probability threshold of 0.48 was selected to maximize F1-score within an 80/20 internal validation split of the development set and was then applied directly to the held-out temporal test.

3.9 Model evaluation

The final model was evaluated on the held-out NHANES 2021–2023 temporal test set. Primary performance metrics included ROC-AUC, area under the precision-recall curve (PR-AUC) [31], F1-score, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Matthews Correlation Coefficient (MCC), and Brier score. Uncertainty around performance estimates was quantified using 1,000 stratified bootstrap resamples with replacement to generate 95% confidence intervals. Calibration was assessed via a calibration plot and numerically using calibration intercept, calibration slope, and expected calibration error (ECE) in 10 bins. Decision curve analysis (DCA) [32] was conducted across threshold probabilities 0.01 to 0.80. Exploratory subgroup performance was evaluated descriptively within the held-out NHANES 2021–2023 temporal test set across sex, age group (12-15 vs 16-19 years), race/ethnicity, and BMI tertile. Because the temporal test set consisted of a single held-out survey period and contained only 37 MetS-positive adolescents, subgroup estimates were interpreted cautiously and were not used for model selection, threshold optimisation, or formal fairness equivalence testing.

3.10 Model interpretation and ablation experiment

SHAP (SHapley Additive exPlanations) [33] values were calculated for all 417 adolescents in the temporal test set to interpret the final LightGBM model. A TreeExplainer was applied to the exact final model using the same pre-processing and prediction workflow as the primary analysis. A probability consistency check confirmed that SHAP refit predictions matched the locked model predictions to a maximum absolute difference of 1.11×10^{-16} confirming that SHAP explains the final model and not a mismatched refit.

Ablation experiments quantified the contribution of individual features and modelling decisions by systematically removing predictors or changing pipeline elements and evaluating PR-AUC on the same temporal test set. Key ablations included removing TG/HDL-C ratio, removing the full lipid domain (TG/HDL-C + total cholesterol), removing BMI, removing age, removing the insulin-resistance domain defined as (HOMA-IR + log-insulin), removing SMOTE, and replacing LightGBM with elastic-net logistic regression. PR-AUC was selected as the primary ablation metric because MetS-positive cases were relatively uncommon in the temporal test set.

3.11 Sensitivity analysis

Two pre-specified sensitivity analyses were conducted. First, the Vitamin D sensitivity analysis evaluated a model augmented with vitamin D (ng/mL) restricted to participants with observed values. This analysis included (training n = 2,465, MetS-positive = 186; test n = 417, MetS-positive = 37) using three approaches: PMM-style imputed, missing-flag indicator, and observed value subset only. Second, a recalibration sensitivity analysis compared the

DOI: 10.5281/zenodo.21992844

primary uncalibrated model against Platt scaling and isotonic regression calibration.

3.12 Software

All analyses were performed in Python 3.11. Key packages included: pandas and numpy for data management; scikit-learn [27] 1.x for Iterative Imputer, Elastic Net, and RFECV; LightGBM [19] 4.x for the primary classifier; imbalanced learn [30] 0.11.x for SMOTE; shap [33] 0.44.x for SHAP interpretation; matplotlib and seaborn for visualisation. Survey-weighted descriptive statistics used fasting subsample weights appropriate for fasting laboratory analyses, following NHANES analytic guidelines for pooled multi-cycle analyses [23].

The final analytic sample contained 3,567 adolescents (Table 1). A total of 281 met the criteria for MetS, corresponding to an unweighted prevalence of 7.88% and a weighted prevalence of 7.34%. The development set consisted of 3,150 adolescents, including 244 MetS-positive participants (7.75%), whereas the temporal test set included 417 adolescents, of whom 37 were MetS-positive (8.87%). Adolescents with MetS had markedly higher BMI (33.25 vs 23.43 kg/m²), waist circumference (106.08 vs 80.52 cm), TG/HDL-C ratio (3.85 vs 1.46), HOMA-IR (7.28 vs. 2.86), and uric acid (6.26 vs 5.03 mg/dL), and substantially lower HDL-C (38.40 vs 53.52 mg/dL; all $p < 0.001$). Male sex (67.4% vs 50.4%) and Mexican American ethnicity (22.0% vs 13.6%) were significantly more prevalent in the MetS group.

4. Results

4.1 Study population and baseline characteristics

Table 1. Weighted Baseline Characteristics of analytic sample by MetS status

(NHANES 2007-2023)

Characteristic	Total N = 3,567	Non-MetS N = 3,286	MetS N = 281	P value
Age, years	15.48 ± 0.05	15.49 ± 0.05	15.35 ± 0.18	0.145
BMI, kg/m²	24.16 ± 0.14	23.43 ± 0.13	33.25 ± 0.53	<0.001
Waist circumference, cm	82.40 ± 0.35	80.52 ± 0.32	106.08 ± 1.25	<0.001
Systolic BP, mmHg	108.90 ± 0.23	108.36 ± 0.23	115.81 ± 0.89	<0.001
Diastolic BP, mmHg	60.88 ± 0.25	60.54 ± 0.26	65.13 ± 1.00	<0.001
Triglycerides, mg/dL	77.44 ± 1.03	72.30 ± 0.94	142.20 ± 4.74	<0.001
HDL-C, mg/dL	52.41 ± 0.27	53.52 ± 0.27	38.40 ± 0.55	<0.001
Fasting glucose, mg/dL	95.71 ± 0.34	95.23 ± 0.35	101.68 ± 1.20	<0.001
Insulin, µIU/mL	13.29 ± 0.24	12.07 ± 0.20	28.64 ± 1.48	<0.001
Log insulin	2.38 ± 0.01	2.32 ± 0.01	3.18 ± 0.05	<0.001
HOMA-IR	3.19 ± 0.06	2.86 ± 0.05	7.28 ± 0.42	<0.001
LDL-C, mg/dL	87.92 ± 0.59	87.05 ± 0.60	98.94 ± 2.53	<0.001
Total cholesterol, mg/dL	155.76 ± 0.67	154.96 ± 0.69	165.73 ± 2.79	<0.001
Non-HDL-C, mg/dL	103.35 ± 0.67	101.45 ± 0.66	127.33 ± 2.68	<0.001
TyG index	8.08 ± 0.01	8.02 ± 0.01	8.78 ± 0.04	<0.001
TG/HDL-C ratio	1.64 ± 0.03	1.46 ± 0.03	3.85 ± 0.15	<0.001
Uric acid, mg/dL	5.12 ± 0.03	5.03 ± 0.03	6.26 ± 0.10	<0.001
ALT/GGT ratio	1.38 ± 0.02	1.38 ± 0.02	1.37 ± 0.05	0.650
Vitamin D, ng/mL	25.36 ± 0.23	25.59 ± 0.24	22.37 ± 0.59	<0.001
Log CRP	-0.51 ± 0.04	-0.60 ± 0.04	0.62 ± 0.10	<0.001
Sex	—	—	—	**<0.001**
Male	51.7%	50.4%	67.4%	—
Female	48.3%	49.6%	32.6%	—
Race/ethnicity	—	—	—	**<0.001**
Mexican American	14.2%	13.6%	22.0%	—

DOI: 10.5281/zenodo.21992844

Non-Hispanic Black	14.0%	14.3%	10.0%	—
Non-Hispanic White	54.9%	55.2%	51.0%	—
Other / Asian / Multi-Racial	9.4%	9.6%	7.3%	—
Other Hispanic	7.5%	7.3%	9.6%	—
NHANES cycle	—	—	—	**0.268**
2007-2008	13.8%	13.7%	14.3%	—
2009-2010	12.9%	13.2%	9.6%	—
2011-2012	13.5%	13.7%	10.6%	—
2013-2014	13.3%	13.2%	15.2%	—
2015-2016	13.3%	13.2%	14.3%	—
2017-March 2020	19.8%	19.7%	21.6%	—
2021-2023	13.4%	13.3%	14.3%	—

Values are weighted mean ±SE for continuous variables and weighted percentages for categorical variables. P values are from group comparison tests; indicates not applicable.

- BMI = body mass index; BP = blood pressure;
- HDL-C = high-density lipoprotein cholesterol;
- HOMA-IR = Homeostatic Model Assessment for Insulin Resistance;
- TG = triglycerides; TyG = triglyceride-glucose index;
- CRP = C-reactive protein.

Table 2. Training vs temporal test-set characteristics

Characteristic	Training set	Temporal test set	P value
Sample size, n	3,150	417	—
MetS-positive, n (%)	244 (7.7%)	37 (8.9%)	—
Age, years	15.48 ± 2.26	15.53 ± 2.23	0.662
BMI, kg/m²	24.32 ± 6.32	24.73 ± 6.55	0.229
Waist circumference, cm	82.41 ± 15.57	83.54 ± 15.51	0.163
Systolic BP, mmHg	109.42 ± 10.10	107.00 ± 9.97	<0.001
Diastolic BP, mmHg	59.98 ± 11.95	64.54 ± 7.70	<0.001
Triglycerides, mg/dL	76.90 ± 47.65	84.27 ± 45.39	0.002
HDL-C, mg/dL	52.85 ± 12.18	51.74 ± 10.66	0.050
Fasting glucose, mg/dL	95.83 ± 17.50	94.85 ± 7.16	0.036
HOMA-IR	3.60 ± 3.83	4.20 ± 10.34	0.238
Log insulin	2.47 ± 0.66	2.50 ± 0.72	0.294
Uric acid, mg/dL	5.08 ± 1.25	5.03 ± 1.35	0.407
Total cholesterol, mg/dL	155.80 ± 29.65	154.28 ± 29.41	0.321
TG/HDL-C ratio	1.63 ± 1.37	1.77 ± 1.26	0.034
Sex	—	—	**0.015**
Male	1,669 (53.0%)	194 (46.5%)	—
Female	1,481 (47.0%)	223 (53.5%)	—
Race/ethnicity	—	—	**<0.001**
Mexican American	697 (22.1%)	67 (16.1%)	—
Non-Hispanic Black	793 (25.2%)	51 (12.2%)	—
Non-Hispanic White	926 (29.4%)	171 (41.0%)	—

DOI: 10.5281/zenodo.21992844

Other / Asian / Multi-Racial	400 (12.7%)	62 (14.9%)	—
Other Hispanic	334 (10.6%)	66 (15.8%)	—

Values are mean ± SD for continuous variables and n (%) for categorical variables. P values are from group comparison tests; — indicates not applicable.. BP = blood pressure; HDL-C = high-density lipoprotein cholesterol; HOMA-IR = Homeostatic Model Assessment for Insulin Resistance; TG = triglycerides.

4.2 MetS component profile

Component-level analysis indicated that MetS-positive adolescents had the expected clustering of central adiposity, dyslipidaemia, and impaired glycaemic risk. High waist circumference using age- and sex-specific p90 criteria was present in 702 participants (19.68%), elevated triglycerides ≥110 mg/dL in 622 (17.44%), low HDL-C ≤40 mg/dL in 504 (14.13%) and elevated fasting glucose ≥100 mg/dL in 896 (25.12%). Elevated blood pressure was the least frequent component, present in 234 participants (6.56%), consistent with prior NHANES-based findings on adolescent MetS [34]. Although the waist criterion was based on an age- and sex-specific 90th percentile reference, 19.68% of the analytic sample exceeded this threshold because the reference values were external paediatric cut-offs rather than percentiles recalculated within the present NHANES sample. The relatively high prevalence of elevated fasting glucose reflects the fixed ≥100 mg/dL threshold used in the Johnson/Ford modified ATP III definition. [6] A diagnostic definition sensitivity analysis showed that an adult AHA/NHLBI-like waist-cut-off definition classified 223 of 3,567 adolescents as MetS-positive, corresponding to a prevalence of 6.25%, compared with 281 of 3,567 adolescents, or 7.88%, under the Johnson/Ford paediatric definition. Thus, adult waist cut-offs identified 58 fewer MetS-positive adolescents, representing 20.6% fewer cases than the paediatric definition. The distribution of individual MetS component positivity by MetS status is shown in Figure 2.

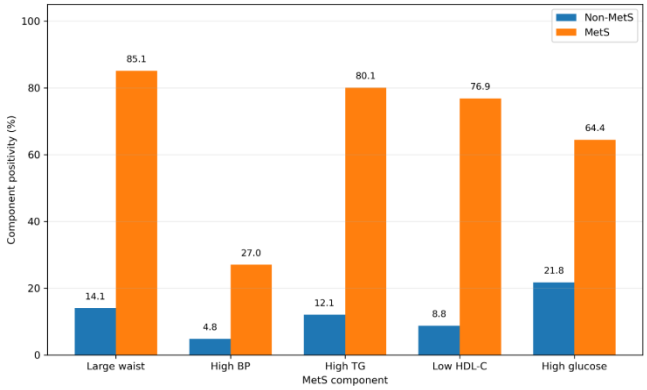

Figure 2. Prevalence of individual metabolic syndrome (MetS) components among adolescents with and without MetS. Component positivity was substantially higher in the MetS group across all five diagnostic components, particularly large waist circumference, elevated triglycerides, and low HDL-C.

4.3 Primary temporal test performance

The final LightGBM + SMOTE model demonstrated strong temporal discrimination on the held-out NHANES 2021–2023 test set (Figure 3; Table 3, Panel A). Using the pre-specified threshold of 0.48, the model correctly detected 29 of 37 MetS-positive adolescents and generated 11 false positives among 380 MetS-negative adolescents, corresponding to (TP/FP/FN/TN = 29/11/8/369). The model achieved ROC-AUC 0.977 (95% CI: 0.959–0.991), PR-AUC 0.868 (0.774–0.943), and NPV 0.979 (0.963–0.992), indicating that a negative model prediction reliably excluded MetS in approximately 98% of adolescents screened. The Brier score was 0.031 (0.021–0.043), indicating well-calibrated predicted probabilities. Calibration should be interpreted descriptively given the small number of test set positives (n = 37). Decision curve analysis confirmed net clinical benefit over treat-all and treat-none strategies across threshold probabilities from 5% to 50%. The Brier score indicated low overall prediction error. Additional calibration metrics showed an observed/expected ratio of 0.791, calibration intercept of -0.610, calibration slope of 1.156, and 10-bin expected calibration error of 0.027. These estimates suggest possible overprediction of absolute risk and should be interpreted descriptively because the temporal test set included only 37 MetS-positive adolescents.

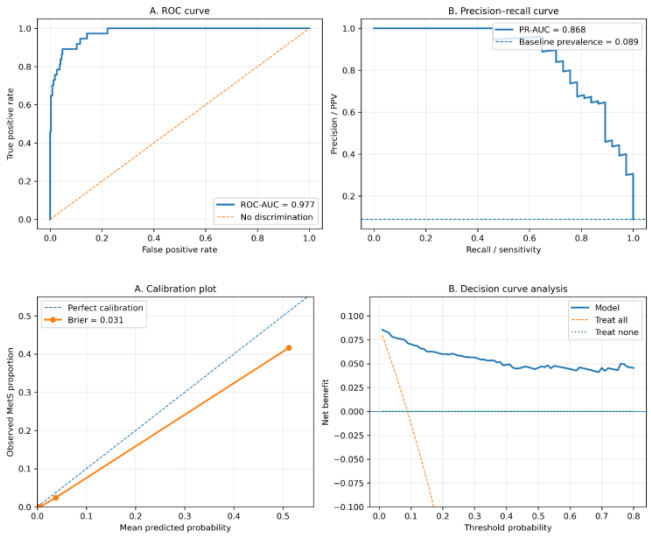

Figure 3. Model discrimination, calibration, and clinical utility in the held-out NHANES 2021–2023 temporal test set. Performance was evaluated using ROC and precision–recall curves, calibration analysis, and decision curve analysis.

Table 3. Primary temporal test performance and key ablation analysis in the NHANES 2021–2023 temporal test set; n = 417; MetS-positive = 37.

Panel A. Primary model performance

Metric	Point estimate	95% Bootstrap CI (1,000 resamples)
ROC-AUC	0.977	0.959 – 0.991
PR-AUC	0.868	0.774 – 0.943
F1-score	0.753	0.633 – 0.854
Sensitivity	0.784	0.641 – 0.914
Specificity	0.971	0.953 – 0.987
PPV	0.725	0.590 – 0.862
NPV	0.979	0.963 – 0.992
MCC	0.729	0.605 – 0.840

DOI: 10.5281/zenodo.21992844

Brier score	0.031	0.021 – 0.043
Decision threshold	0.48	—
TP / FP / FN / TN	29 / 11 / 8 / 369	—

Panel A. Primary model performance on the held-out NHANES 2021–2023 temporal test set. Point estimates are reported with 95% bootstrap confidence intervals based on 1,000 resamples. Performance at the selected decision threshold (0.48) is summarized by sensitivity, specificity, PPV, NPV, F1-score, MCC, and TP/FP/FN/TN counts.

Panel B. Key ablation experiments using PR-AUC on the temporal test set

Ablation experiment	PR-AUC	ΔPR-AUC vs refit reference
Refit 8-variable reference model	0.874	—
Remove TG/HDL-C ratio	0.507	−0.367
Remove lipid domain (TG/HDL-C + total cholesterol)	0.500	−0.373
Remove BMI	0.851	−0.023
Remove age	0.777	−0.096
Remove insulin-resistance domain (HOMA-IR + log-insulin)	0.868	−0.006
No SMOTE (no imbalance adjustment)	0.838	−0.036
Elastic-net logistic regression + SMOTE	0.802	−0.072

Panel B. Refit ablation comparisons on the same temporal test set. The reference model in Panel B was refit under the ablation-analysis protocol; therefore, its PR-AUC differs slightly from the primary model reported in Panel A. Values are intended for relative comparison of feature domains and modelling choices, not as independent primary performance estimates.

4.4 Model Interpretation and ablation experiments

TyG index and ALT/GGT ratio were not included in the final 8-variable model, indicating that these two variables did not show any additional incremental value in prediction beyond the combination of TG/HDL-C ratio, BMI, HOMA-IR, log-insulin, uric acid, total cholesterol, age, and sex. SHAP analysis (Figure 4A) showed that TG/HDL-C ratio was the dominant predictor in the final model, with the highest mean |SHAP| value (1.924), followed by BMI (1.236) and HOMA-IR (1.160), collectively reflecting the dyslipidaemia and insulin resistance axis central to MetS pathophysiology. Age ranked fourth (mean |SHAP| = 0.854) with an inverse association with predicted risk, suggesting older adolescents had lower predicted risk after conditioning on other predictors. Sex also contributed to model predictions in the SHAP analysis, although this demographic association should be interpreted descriptively rather than causally. Uric acid contributed positively (mean |SHAP| = 0.455), consistent with its known association with insulin resistance and hyperuricaemia in adolescent metabolic disease. [35] Total cholesterol (0.316) and log-insulin (0.300) had smaller individual contributions, partly reflecting conditional suppression by TG/HDL-C and HOMA-IR within the tree ensemble.

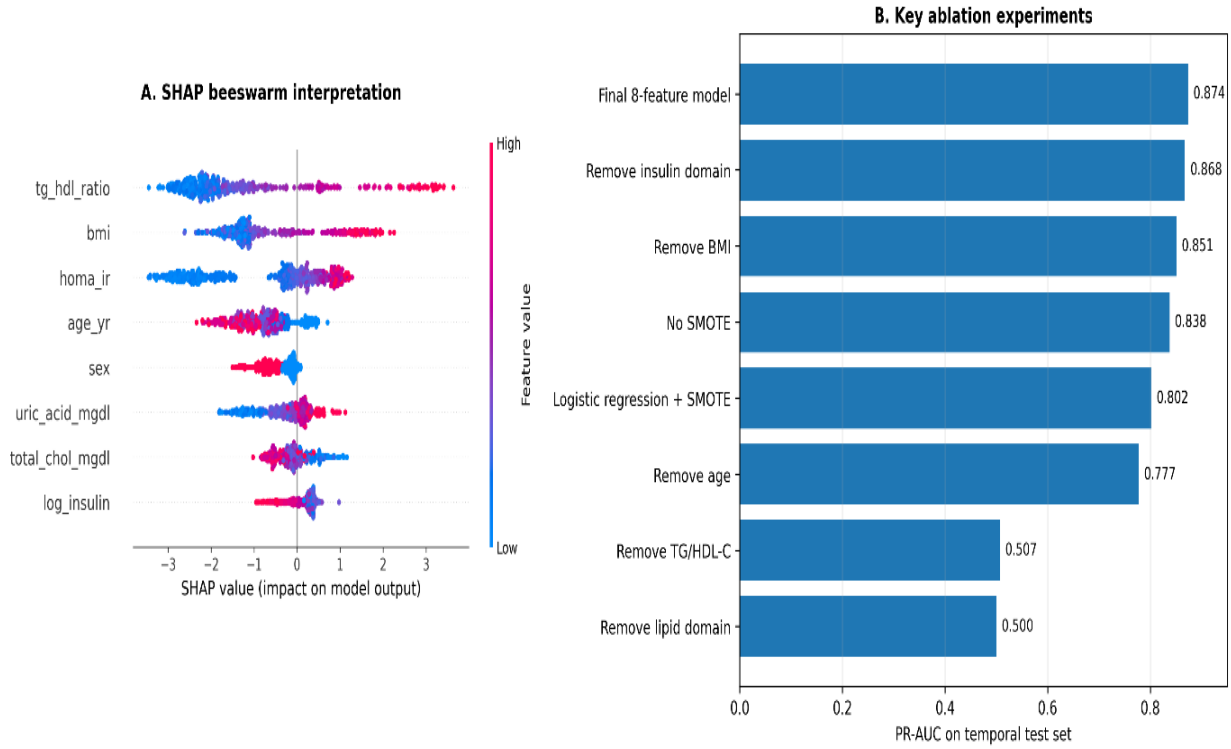

Figure 4. Model interpretation and ablation analysis. (Panel A) shows the SHAP beeswarm plot illustrating the direction and magnitude of each predictor's contribution to the final LightGBM + SMOTE model. (Panel B) shows PR-AUC under key ablation experiments on the temporal test set.

DOI: 10.5281/zenodo.21992844

Ablation experiments (Figure 4B) confirmed that the lipid domain carried the largest incremental predictive signal. Removing the TG/HDL-C ratio reduced PR-AUC from 0.874 to 0.507 (ΔPR-AUC = −0.367), and removing the full lipid domain reduced PR-AUC to 0.500 (−0.373), indicating a substantial reduction in predictive performance when lipid-derived information was removed. Removing BMI produced a smaller reduction (PR-AUC 0.851; −0.023), and removing age reduced PR-AUC to 0.777 (−0.096). Notably, removing the entire insulin-resistance domain (HOMA-IR and log-insulin) produced only a marginal reduction (PR-AUC 0.868; −0.006), suggesting that the lipid and adiposity features retained overlapping insulin resistance signal. SMOTE improved performance relative to no imbalance adjustment (0.874 vs 0.838), and LightGBM outperformed elastic net logistic regression (0.874 vs 0.802), supporting the final model architecture.

4.5 Vitamin D sensitivity

Among participants with available vitamin D values, adding vitamin D to the primary model did not improve temporal test PR-AUC performance under any of the three approaches: The PMM-style imputed model achieved a PR-AUCof 0.855, the missing flag model (0.852), and the observed subset model (0.832), all below the primary model without vitamin D (PR-AUC 0.868). These findings suggest that vitamin D did not add meaningful incremental predictive value in this dataset. Vitamin D levels differed significantly between adolescents with and without MetS. (MetS 22.37 vs non-MetS 25.59 ng/mL; $p<0.001$), this difference is captured by the retained metabolic predictors after conditioning within the ensemble model.

4.6 Recalibration sensitivity:

Post-hoc recalibration was evaluated using Platt scaling and isotonic regression. Neither method improved temporal test PR-AUC compared with the primary uncalibrated model (PR-AUC 0.868). Therefore, the raw LightGBM + SMOTE model trained on the full development set was retained for primary reporting because it achieved the best PR-AUC among the recalibration variants.

4.7 Exploratory subgroup performance

Exploratory subgroup performance was assessed within the held-out NHANES 2021–2023 temporal test set across sex, age group, race/ethnicity, and BMI tertile. Because only 37 MetS-positive adolescents were present in the temporal test set, these analyses were considered descriptive and were not used for model selection, threshold optimisation, or formal fairness/equivalence testing. Discrimination remained high by ROC-AUC across sex strata, although PR-AUC was lower among females than males: ROC-AUC 0.961 and PR-AUC 0.702 among females, compared with ROC-AUC 0.985 and PR-AUC 0.933 among males. Age-stratified performance was similar, with PR-AUC 0.890 among adolescents aged 12–15 years and 0.876 among those aged 16–19 years. Race/ethnicity and BMI-tertile estimates were more unstable because several strata contained fewer than 10 MetS-positive adolescents; for example, Non-Hispanic Black adolescents included only one positive case, while the first and second BMI tertiles included two and four positive cases, respectively. The highest BMI tertile, which contained 31 MetS-positive adolescents, showed PR-AUC 0.925. No consistent subgroup-specific failure pattern was identified, but these findings should be interpreted as exploratory and underpowered.

5. Discussion

This study created and temporally validated a laboratory-enhanced machine learning model for predicting metabolic syndrome (MetS) in U.S. adolescents using seven NHANES survey periods spanning 2007–2023, the longest data window applied to this research question to date. The final 8-variable LightGBM + SMOTE model achieved ROC-AUC 0.977 (95% CI: 0.959–0.991) and PR-AUC 0.868 (0.774-0.943) on the temporally held-out 2021-2023 post-pandemic test cohort, demonstrating that laboratory-enhanced MetS prediction is temporally robust during a 16-year observation window. Three methodological contributions distinguish this work from prior studies: the application of paediatric-appropriate Johnson/Ford waist circumference criteria, the use of a temporal rather than random validation design, and the use of a comprehensive laboratory panel with formal ablation experiments to quantify the contribution of each predictor domain.

5.1 Comparison with prior work

The closest prior study was conducted by Zhang et al. [9], who published a LightGBM-based model for adolescent MetS based on 139 adolescents who were MetS-positive across five cycles of NHANES (2007–2016), with an ROC-AUC of 0.969 [9]. The present study builds on that research in three important ways. First, Zhang et al. used the adult AHA/NHLBI waist circumference thresholds (102 cm for males, 88 cm for females) in a paediatric population. In this analytic sample, the adult AHA/NHLBI-like waist-cut-off definition classified 223 of 3,567 adolescents as MetS-positive, corresponding to a prevalence of 6.25%, whereas the Johnson/Ford paediatric definition classified 281 of 3,567 adolescents as MetS-positive, corresponding to a prevalence of 7.88%. Thus, adult waist cut-offs identified 58 fewer MetS-positive adolescents, or 20.6% fewer cases than the paediatric definition. Second, Zhang et al. used a random 7:3 train-test split of the same five survey cycles. Training and test observations were sampled from the same temporal distribution, so this design is not appropriate for testing the model's ability to generalise to a future sample of observations [36]. The temporal validation approach used here trains on survey periods from 2007–2020 and evaluates performance on the entirely held-out August 2021–August 2023 cycle, collected after NHANES field operations resumed following the COVID-19-related interruption [24], including post-pandemic shifts in adolescent obesity prevalence and healthcare access [8]. The strong performance of the present model on this more challenging test set (ROC-AUC 0.977) provides stronger evidence of temporal generalizability than a random-split design could offer. Third, Zhang et al.[9] excluded all biochemical laboratory indicators by design, restricting predictors to anthropometric measurements and demographic variables. The present study shows that a laboratory panel that included insulin resistance indices (HOMA-IR and log-insulin), lipid metabolism markers (TG/HDL-C ratio and total cholesterol), purine metabolism (uric acid), and adiposity/body composition (BMI) achieves strong temporal discrimination with mechanistically interpretable predictor contributions confirmed by SHAP analysis.

5.2 Interpretation of SHAP and ablation results

SHAP analysis identified TG/HDL-C ratio as the most influential predictor (mean |SHAP| = 1.924), followed by BMI (1.236) and HOMA-IR (1.160), reflecting the dyslipidaemia-adiposity insulin resistance axis central to MetS pathophysiology. Age ranked fourth (mean|SHAP| = 0.854) and showed an inverse association with

DOI: 10.5281/zenodo.21992844

predicted risk after conditioning on other predictors. Sex also contributed to model predictions in the SHAP analysis, although this demographic association should be interpreted descriptively rather than causally. Uric acid contributed positively (mean|SHAP| = 0.455), consistent with its role as a non-lipid metabolic marker. Total cholesterol and log-insulin had smaller individual contributions, likely because part of their predictive information was already captured by TG/HDL-C ratio and HOMA-IR within the tree ensemble.

Ablation experiments supported the SHAP rankings. Removing TG/HDL-C ratio reduced PR-AUC from 0.874 to 0.507 and removing the full lipid domain reduced PR-AUC to 0.500, indicating a substantial reduction in predictive performance when lipid-derived information was removed. Removing the insulin-resistance domain produced only a small reduction in PR-AUC, suggesting that TG/HDL-C ratio and BMI captured overlapping insulin-resistance information within the ensemble. This does not imply that insulin resistance is biologically unimportant; rather, it indicates conditional redundancy among correlated metabolic predictors in the final model.

5.3 Methodological Strength

The choice of SMOTE-only [30] using a minority-to-majority sampling ratio of 0.50 was based on a four-strategy ablation comparison rather than an arbitrary decision, and SMOTE was used only within the training pipeline to avoid the data leakage that can occur when SMOTE is applied before the train-test split. Removing SMOTE reduced PR-AUC from 0.874 to 0.838, and replacing LightGBM with elastic-net logistic regression reduced PR-AUC to 0.802, suggesting that both the non-linear ensemble structure and imbalance correction contributed meaningfully to final performance. Iterative multivariate imputation was also fitted within the training pipeline only, allowing missing predictor values to be handled without using information from the temporal test set.

5.4 Vitamin D

Vitamin D showed a significant bivariate difference between MetS and non-MetS adolescents (22.37 vs. 25.59 ng/mL; $p < 0.001$), yet adding it to the model did not enhance the temporal PR-AUC under any of the sensitivity settings compared to the primary model (0.868): PMM-imputed (0.855), missing-flag (0.852), and observed-subset (0.832). This suggests that once TG/HDL-C ratio, HOMA-IR, and BMI are conditioned on, vitamin D does not add meaningful discriminative information in this dataset. Its structural absence from the 2017–March 2020 NHANES cycle further limits its use as a primary predictor.

5.5 Clinical Implications

An NPV of 0.979 (95% CI: 0.963–0.992) confirms that a negative model prediction reliably excludes MetS in approximately 98% of adolescents, a valuable property for population screening. Decision curve analysis found net clinical benefit between treat-all and treat-none strategies across threshold probabilities of 5%–50%, covering the range relevant to school- and community-based screening. The model is reasonably presented as a screening and risk-ranking tool rather than a causal biomarker discovery model, given that TG/HDL-C and HOMA-IR partially overlap with MetS diagnostic components.

5.6 Limitations

Subgroup robustness analyses were exploratory because the held-out temporal test set contained only 37 MetS-positive adolescents; therefore, subgroup-specific estimates, particularly for race/ethnicity and BMI tertiles, were underpowered and should not be interpreted as definitive evidence of fairness or subgroup equivalence. Adolescents were eligible for the final analytic sample only if complete data were available for all five MetS diagnostic components, which reduced the sample to 3,567 participants from 10,178 eligible adolescents. This attrition may lead to selection bias, so prevalence estimates and model performance may only be valid for the diagnostic-complete fasting analytic sample. The temporal test set contained only 37 MetS-positive cases, creating large confidence intervals for metrics that depend on the threshold. Calibration results should therefore be interpreted descriptively and cautiously. NHANES is a cross-sectional study [23], which limits causal inference and the assessment of longitudinal MetS pathways. In addition, the model has not yet been externally validated in an independent cohort [36, 37], and such validation is required before clinical application.

6. Conclusion

This study created and temporally validated a laboratory-enhanced machine learning model for adolescent metabolic syndrome using seven pooled NHANES survey periods from 2007–2023, the largest observation window available to this question to date. By defining MetS using paediatric-appropriate Johnson/Ford criteria rather than adult anthropometric thresholds, this analysis identified 20.6% more MetS-positive adolescents than would have been identified by an adult-derived waist-circumference definition within the same sample, emphasizing that definitional choice alone materially changes who is classified as at risk in youth populations.

The final eight-variable LightGBM model, trained with SMOTE-based imbalance correction and evaluated on a temporally independent August 2021-August 2023 cohort following the COVID-19-related interruption in NHANES data collection, achieved strong and clinically meaningful discrimination (ROC-AUC 0.977; PR-AUC 0.868) and an NPV of 0.979, confirming that it could be used as a population-level screening tool to effectively rule out MetS. Model development, feature selection, imputation, and threshold optimisation were done only on data before the test period, thus reflecting prospective, out-of-time generalisability, as opposed to in-sample interpolation, which is a more rigorous and clinically relevant test than the random train-test splits that have been used in previous adolescent MetS models.

The interpretation of SHAP values and systematic ablation experiments support a consistent biological narrative. TG/HDL-C ratio, BMI, and HOMA-IR collectively form the dyslipidaemia-adiposity insulin resistance axis that drives predictive performance, with the lipid domain alone accounting for the majority of the model's discriminative signal (removal reduced PR-AUC by 0.373). Notably, removing the insulin-resistance domain produced only a marginal performance loss, indicating that lipid and adiposity markers absorb much of that signal within the ensemble rather than that insulin resistance is biologically irrelevant to MetS pathogenesis. Neither the TyG index nor the ALT/GGT ratio nor supplementary vitamin D measurements added incremental predictive value once the core laboratory panel was accounted for, and post-hoc recalibration did not improve upon the uncalibrated model.

DOI: 10.5281/zenodo.21992844

Together, these findings show that a laboratory-enhanced, paediatric criteria-based machine learning approach can achieve high-level discriminative accuracy during a sixteen-year period that includes a major disruption to adolescent health behaviours and healthcare access, extending prior anthropometry-only models built on adult diagnostic thresholds and random-split validation. The model should be positioned as a risk-stratification and screening aid rather than a causal or diagnostic instrument, given that several of its top predictors mechanistically overlap with the MetS diagnostic components themselves. Given the modest number of MetS-positive cases in the temporal holdout (n=37), the exclusion of adolescents with incomplete fasting laboratory data, and the absence of an external validation cohort, prospective validation in independent, laboratory-complete adolescent samples, ideally with longer-term longitudinal follow-up to assess transition to overt cardiometabolic disease, is warranted before this approach is considered for clinical and school-based screening deployment.

References

1. Morrison, J.A., L.A. Friedman, and C. Gray-McGuire, *Metabolic syndrome in childhood predicts adult cardiovascular disease 25 years later: the Princeton Lipid Research Clinics Follow-up Study.* Pediatrics, 2007. **120**(2): p. 340–345.
2. Morrison, J.A., et al., *Metabolic syndrome in childhood predicts adult metabolic syndrome and type 2 diabetes mellitus 25 to 30 years later.* Journal of Pediatrics, 2008. **152**(2): p. 201–206.
3. Zhang, B., et al., *Metabolic syndrome in children and adolescents: definitions, epidemiology, pathophysiology, interventions, and challenges.* Frontiers in endocrinology, 2025. **16**: p. 1512642.
4. Cleeman, J.I., *Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III).* 2001.
5. Grundy, S.M., et al., *Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement.* Circulation, 2005. **112**(17): p. 2735–2752.
6. Johnson, W.D., et al., *Prevalence of risk factors for metabolic syndrome in adolescents: National Health and Nutrition Examination Survey (NHANES), 2001-2006.* Archives of pediatrics & adolescent medicine, 2009. **163**(4): p. 371–377.
7. Li, C., et al., *Recent trends in waist circumference and waist-height ratio among US children and adolescents.* Pediatrics, 2006. **118**(5): p. e1390–e1398.
8. Krause, K.H., *Disruptions to school and home life among high school students during the COVID-19 pandemic—Adolescent Behaviors and Experiences Survey, United States, January–June 2021.* MMWR supplements, 2022. **71**.
9. Zhang, Y.-z., et al., *Machine Learning-Based predictive model for adolescent metabolic syndrome: Utilizing data from NHANES 2007–2016.* Scientific reports, 2025. **15**(1): p. 3274.
10. Association, A.H., *Diagnosis and management of the metabolic syndrome. An American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Executive summary.* Cardiol Rev, 2005. **13**(6): p. 322–327.
11. Zimmet, P., K.G.M. Alberti, and M.S. Ríos, *A new International Diabetes Federation (IDF) worldwide definition of the metabolic syndrome: the rationale and the results.* 2005, Elsevier Doyma. p. 1371–1375.
12. Cook, S., et al., *Prevalence of a metabolic syndrome phenotype in adolescents: findings from the third National Health and Nutrition Examination Survey, 1988-1994.* Archives of pediatrics & adolescent medicine, 2003. **157**(8): p. 821–827.
13. de Ferranti, S.D., et al., *Prevalence of the metabolic syndrome in American adolescents: findings from the Third National Health and Nutrition Examination Survey.* Circulation, 2004. **110**(16): p. 2494–2497.
14. Weiss, R., et al., *Obesity and the metabolic syndrome in children and adolescents.* New England journal of medicine, 2004. **350**(23): p. 2362–2374.
15. Alberti, G., P. Bennett, and J. Shaw, *The metabolic syndrome in children and adolescents? an IDF consensus report.* Pediatric diabetes, 2007.
16. Zimmet, P., et al., *The metabolic syndrome in children and adolescents.* The Lancet, 2007. **369**(9579): p. 2059.
17. Cook, S., et al., *Metabolic syndrome rates in united states adolescents, from the national health and nutrition examination survey, 1999–2002.* The Journal of pediatrics, 2008. **152**(2): p. 165–170. e2.
18. Duncan, G.E., S.M. Li, and X.-H. Zhou, *Prevalence and trends of a metabolic syndrome phenotype among US adolescents, 1999–2000.* Diabetes care, 2004. **27**(10): p. 2438–2443.
19. Ke, G., et al., *Lightgbm: A highly efficient gradient boosting decision tree.* Advances in neural information processing systems, 2017. **30**.
20. Simental-Mendía, L.E., M. Rodríguez-Morán, and F. Guerrero-Romero, *The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects.* Metabolic syndrome and related disorders, 2008. **6**(4): p. 299–304.
21. Giannini, C., et al., *The triglyceride-to-HDL cholesterol ratio: association with insulin resistance in obese youths of different ethnic backgrounds.* Diabetes care, 2011. **34**(8): p. 1869–1874.
22. Zhang, C., et al., *From traditional metabolic markers to ensemble learning: comparative application of machine learning models for predicting NAFLD risk in adolescents.* Frontiers in Endocrinology, 2025. **16**: p. 1681686.

DOI: 10.5281/zenodo.21992844

23. Statistics, N.C.f.H., *NHANES survey methods and analytic guidelines. Centers for Disease Control and Prevention*. 2024.

24. Terry, A., et al., *Plan and operations of the National Health and Nutrition Examination Survey, August 2021–August 2023*. 2024.

25. Matthews, D.R., et al., *Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man*. diabetologia, 1985. **28**(7): p. 412–419.

26. Program, N.H.B.P.E., *The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents*. 2005: US Department of Health and Human Services, National Institutes of Health ….

27. Pedregosa, F., et al., *Scikit-learn: Machine learning in Python.* the Journal of machine Learning research, 2011. **12**: p. 2825–2830.

28. Zou, H. and T. Hastie, *Regularization and variable selection via the elastic net.* Journal of the Royal Statistical Society Series B: Statistical Methodology, 2005. **67**(2): p. 301–320.

29. Guyon, I., et al., *Gene selection for cancer classification using support vector machines.* Machine learning, 2002. **46**(1): p. 389–422.

30. Chawla, N.V., et al., *SMOTE: synthetic minority over-sampling technique.* Journal of artificial intelligence research, 2002. **16**: p. 321–357.

31. Saito, T. and M. Rehmsmeier, *The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets.* PloS one, 2015. **10**(3): p. e0118432.

32. Vickers, A.J. and E.B. Elkin, *Decision curve analysis: a novel method for evaluating prediction models.* Medical Decision Making, 2006. **26**(6): p. 565–574.

33. Lundberg, S.M. and S.-I. Lee, *A unified approach to interpreting model predictions.* Advances in neural information processing systems, 2017. **30**.

34. Miller, J.M., et al., *Prevalence of metabolic syndrome and individual criterion in US adolescents: 2001–2010 National Health and Nutrition Examination Survey.* Metabolic syndrome and related disorders, 2014. **12**(10): p. 527–532.

35. DeBoer, M.D., L. Dong, and M.J. Gurka, *Racial/ethnic and sex differences in the relationship between uric acid and metabolic syndrome in adolescents: an analysis of National Health and Nutrition Survey 1999-2006.* Metabolism, 2012. **61**(4): p. 554–561.

36. Steyerberg, E.W. and F.E. Harrell Jr, *Prediction models need appropriate internal, internal-external, and external validation.* Journal of clinical epidemiology, 2015. **69**: p. 245.

37. Collins, G.S., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.* Journal of British Surgery, 2015. **102**(3): p. 148–158.

DOI: 10.5281/zenodo.21992844