



OPEN ACCESS

EDITED BY

George Grant,
Independent Researcher, Aberdeen,
United Kingdom

REVIEWED BY

Giovanni Tarantino,
University of Naples Federico II, Italy
Afshan Masood,
King Saud University Medical City,
Saudi Arabia
Szymon Suwala,
Nicolaus Copernicus University in Toruń,
Poland

*CORRESPONDENCE

Weining Xie
✉ xwn1219@qq.com

RECEIVED 22 April 2025

ACCEPTED 09 June 2025

PUBLISHED 30 June 2025

CITATION

Hong Y, Chen X, Wang L, Zhang F, Zeng Z and
Xie W (2025) Machine learning prediction of
metabolic dysfunction-associated fatty liver
disease risk in American adults using body
composition: explainable analysis based on
SHapley Additive exPlanations.
Front. Nutr. 12:1616229.
doi: 10.3389/fnut.2025.1616229

COPYRIGHT

© 2025 Hong, Chen, Wang, Zhang, Zeng and
Xie. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Machine learning prediction of metabolic dysfunction-associated fatty liver disease risk in American adults using body composition: explainable analysis based on SHapley Additive exPlanations

Yan Hong¹, Xinrong Chen², Ling Wang³, Fan Zhang¹,
ZiYing Zeng¹ and Weining Xie^{4*}

¹Affiliated Guangdong Hospital of Integrated Traditional Chinese and Western Medicine of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine, Foshan, China, ²First Clinical Medical College, Guangzhou University of Chinese Medicine, Guangzhou, China, ³First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China, ⁴Infectious Disease Department, Guangdong Provincial Hospital of Integrated Traditional Chinese and Western Medicine, Foshan, China

Background: Metabolic dysfunction-associated fatty liver disease (MAFLD) is a prevalent and progressive liver disorder closely linked to obesity and metabolic dysregulation. Traditional anthropometric measures such as body mass index (BMI) are limited in their ability to capture fat distribution and associated risk. This study aimed to develop and validate machine learning (ML) models for predicting MAFLD using detailed body composition metrics and to explore the relative contributions of adipose tissue features through explainable ML techniques.

Methods: Data from the 2017–2018 National Health and Nutrition Examination Survey (NHANES) were used to construct predictive models based on anthropometric, demographic, lifestyle, and clinical variables. Six ML algorithms were implemented: decision tree (DT), support vector machine (SVM), generalized linear model (GLM), gradient boosting machine (GBM), random forest (RF), and XGBoost. The Boruta algorithm was used for feature selection, and model performance was evaluated using cross-validation and a validation set. SHapley Additive exPlanations (SHAP) were employed to interpret feature contributions.

Results: Among the six models, the GBM algorithm exhibited the best performance, achieving area under the receiver operating characteristic curve (AUC) values of 0.875 (training) and 0.879 (validation), with minimal fluctuations in sensitivity and specificity. SHAP analysis identified visceral adipose tissue (VAT), BMI, and subcutaneous adipose tissue (SAT) as the most influential predictors. VAT had the highest SHAP value, underscoring its central role in MAFLD pathogenesis.

Conclusion: This study demonstrates the effectiveness of integrating body composition features with machine learning techniques for MAFLD risk prediction. The GBM model offers robust predictive accuracy and interpretability, with potential applications in clinical decision-making and public health screening strategies. SHAP analysis provides meaningful insights into the relative

importance of adiposity measures, reinforcing the value of fat distribution metrics beyond conventional obesity indices.

KEYWORDS

MAFLD, body composition, machine learning, SHAP, NHANES

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a chronic and progressive liver disorder that develops in genetically susceptible individuals in the context of nutritional excess and insulin resistance (IR). The disease spectrum ranges from simple steatosis (nonalcoholic fatty liver, NAFL) to nonalcoholic steatohepatitis (NASH), and may progress to advanced stages such as fibrosis and cirrhosis (1). With the discovery of a strong relationship between NAFLD and metabolic risk factors, it has been renamed in recent years as metabolic dysfunction-associated fatty liver disease (MAFLD) and metabolism-associated steatosis liver disease (MASLD). MASLD is defined as hepatic steatosis accompanied by cardiometabolic abnormalities, in the absence of other causes of steatosis or excessive alcohol consumption (≥ 30 g/day for men and ≥ 20 g/day for women) (2). By contrast, the 2020 diagnostic criteria for MAFLD (3–6), focus more on metabolic abnormalities than alcohol intake (7). Recent meta-analyses have estimated the global prevalence of MAFLD to be as high as 38.77%, which significantly exceeds the prevalence reported under the previous NAFLD criteria (8–10).

MAFLD is strongly associated with an increased risk of atherosclerotic cardiovascular disease (CVD), chronic kidney disease (CKD), hepatic decompensation, and hepatocellular carcinoma (HCC) (11, 12). Emerging evidence suggests that the “liver–spleen axis” plays a critical role in the pathogenesis and progression of MAFLD. Splenomegaly has been positively correlated with central obesity and the severity of hepatic steatosis (13, 14). Animal studies have shown that high-fat diets induce splenic sinusoidal dilation and lipid accumulation in mice, whereas splenectomy significantly increases hepatic immune cell infiltration and the expression of proinflammatory cytokines such as IL-6 and TNF- α (15, 16). These findings suggest that the spleen may play a protective role in metabolic regulation by maintaining immune homeostasis and attenuating excessive inflammatory responses. Moreover, MAFLD is strongly associated with obesity-related chronic inflammation (17). Dysfunctional adipose tissue promotes the release of free fatty acids (FFAs), which exacerbate hepatic steatosis by inducing inflammation and promoting the development of IR (18).

In clinical practice, BMI is widely used to assess general obesity due to its simplicity (19, 20). However, BMI cannot differentiate between fat mass and lean mass, nor does it account for the spatial distribution of adipose tissue (21, 22). Studies have demonstrated that obesity-related metabolic disturbances are closely associated with fat distribution patterns, particularly the accumulation of visceral adipose tissue (VAT) (23–26). Total abdominal fat area (TAFA) has been identified as an independent risk factor, exhibiting stronger associations with CVD, metabolic disorders, and all-cause mortality than BMI (26–29). TAFA is composed of both subcutaneous adipose tissue (SAT) and VAT. SAT can expand physiologically to buffer against ectopic lipid deposition; however, its compensatory capacity may be constrained by genetic predisposition or impaired

adipogenesis. Persistent caloric excess results in pathological accumulation of VAT (30, 31). VAT is regarded as a hallmark of metabolically unhealthy obesity and is independently associated with a wide range of metabolic disturbances (32–34). Its abnormal expansion is indicative of ectopic lipid deposition (35, 36). The visceral-to-subcutaneous fat ratio (VSR), a novel adiposity metric, has been strongly associated with elevated levels of proinflammatory cytokines and the progression of hepatic steatosis (37).

As a subfield of artificial intelligence, machine learning (ML) excels at identifying complex nonlinear relationships within high-dimensional datasets and has shown considerable advantages in disease screening and risk assessment (38, 39). Unlike traditional statistical methods, ML does not require assumptions about variable distributions and is well-suited to capturing intricate interactions and nonlinear associations. Although ML has been increasingly applied in the diagnosis of liver diseases (22, 40, 41), its utility in exploring associations between multidimensional obesity indices and MAFLD remains underinvestigated. In this study, we leveraged data from the National Health and Nutrition Examination Survey (NHANES) to identify obesity-related indices strongly associated with MAFLD using ML techniques. Furthermore, we employed SHapley Additive exPlanations (SHAP) to interpret the contribution of individual features and to develop an interpretable predictive model.

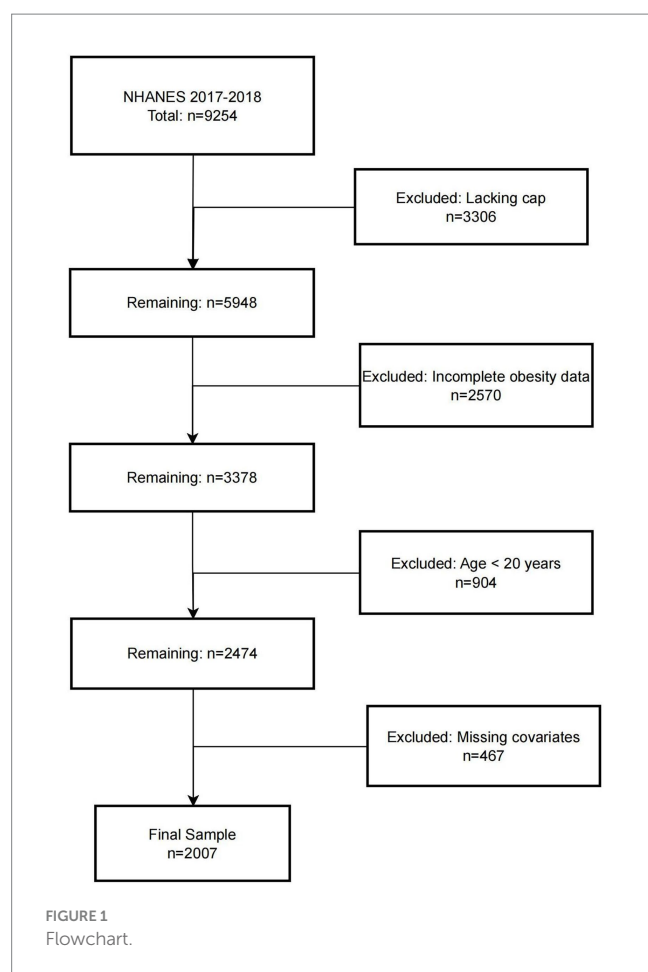
Methods

Participants

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative program jointly conducted by the Centers for Disease Control and Prevention (CDC) and the National Center for Health Statistics (NCHS). The study protocol was approved by the NCHS Research Ethics Review Board, and written informed consent was obtained from all participants. In this study, data from the 2017–2018 NHANES cycle were analyzed. The initial cohort consisted of 9,254 participants. Individuals were sequentially excluded based on the following criteria: (1) lack of hepatic steatosis assessment ($n = 3,306$); (2) missing obesity-related measurements ($n = 2,570$); (3) age < 20 years ($n = 904$); and (4) incomplete covariate data ($n = 467$). A total of 2,007 participants were ultimately included in the final analysis. The screening and selection process is presented in Figure 1.

Definition of MAFLD

MAFLD was diagnosed according to the international expert consensus criteria established in 2020 (7). Hepatic steatosis was assessed using the controlled attenuation parameter (CAP) measured by the FibroScan® 502 V2 Touch device, with a CAP value



≥ 274 dB/m considered indicative of hepatic steatosis (42). In addition, diagnosis required the presence of at least one of the following three conditions:

1. Overweight or obesity: defined as a BMI ≥ 25 kg/m² for Caucasian individuals or ≥ 23 kg/m² for Asian individuals.
2. Type 2 diabetes mellitus (T2DM): diagnosed based on any of the following criteria: (a) fasting plasma glucose (FPG) ≥ 7.0 mmol/L; (b) glycated hemoglobin (HbA1c) $\geq 6.5\%$; (c) a clinical diagnosis of diabetes by a qualified physician.
3. Metabolic dysregulation: defined as the presence of at least two of the following seven criteria:
 - (1) Waist circumference ≥ 102 cm in men or ≥ 88 cm in women (or ≥ 90 cm in Asian men or ≥ 80 cm in Asian women).
 - (2) Blood pressure $\geq 130/85$ mmHg or current use of antihypertensive medication.
 - (3) Plasma triglycerides ≥ 150 mg/dL or treatment with lipid-lowering agents.
 - (4) Plasma high-density lipoprotein (HDL) cholesterol < 40 mg/dL in men or < 50 mg/dL in women, or use of lipid-modifying therapy.
 - (5) Prediabetes (FPG 5.6–6.9 mmol/L; 2-h post-load glucose 7.8–11.0 mmol/L; or HbA1c 5.7–6.4%).
 - (6) Homeostasis model assessment of insulin resistance (HOMA-IR) ≥ 2.5 .

- (7) High-sensitivity C-reactive protein (hs-CRP) > 2 mg/L (7).

Definition of body composition

Anthropometric measurements were conducted by trained NHANES personnel at Mobile Examination Centers (MEC) following standardized protocols. TFAFA(g), VAT(g), and SAT(g) were measured using dual-energy X-ray absorptiometry (DXA), with data automatically processed by Hologic APEX software. The VSR was calculated as the ratio of VAT to SAT. BMI was assessed by trained staff using calibrated instruments to measure height and weight, and was calculated using the formula: weight (kg) / height² (m²).

Covariates

Data extracted from the NHANES database included the following covariates: (1) Demographic characteristics: age, gender (male or female), educational attainment (less than 9th grade; 9th–11th grade; high school graduate or GED equivalent; some college or associate degree; college graduate or above), race (Mexican American; other Hispanic; non-Hispanic White; non-Hispanic Black; other race, including multiracial), marital status (married; widowed; divorced; separated; never married; living with a partner), and family income–poverty ratio (PIR). (2) Lifestyle factors: Alcohol consumption was defined as the intake of more than two standard drinks per day over the past 12 months. Smoking status was defined as having smoked at least 100 cigarettes in a lifetime. (3) Laboratory and clinical measures: These included HbA1c, FPG, HDL, and low-density lipoprotein (LDL). Hypertension was defined as having a systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg based on three separate readings, or a clinical diagnosis of hypertension. Diabetes mellitus was defined as HbA1c $> 6.5\%$ or FPG > 7.0 mmol/L.

Statistical analysis

All statistical analyses were performed using R software (version 4.3.2) and EmpowerStats. Baseline characteristics were summarized according to MAFLD status. Continuous variables were expressed as mean \pm standard deviation (SD), while categorical variables were presented as counts with corresponding percentages. Data visualization was conducted using the ggplot2 package, generating bar plots for categorical variables and histograms for continuous variables. A Pearson correlation matrix was constructed to illustrate inter-variable correlations. To assess multicollinearity, the variance inflation factor (VIF) was calculated through iterative regression modeling. Variables with VIF values exceeding 10 were excluded from subsequent analyses. Feature selection was conducted using the Boruta algorithm, which utilizes shadow features based on random forests to identify the most relevant predictors. A Z-score boxplot was used to visualize feature importance, and the top 10 features most strongly associated with MAFLD were retained for modeling. Prior to model construction, the dataset was randomly partitioned into a training set (70%) and a validation set (30%). Six ML models were developed using the caret package: decision tree (DT), support vector

machine (SVM), generalized linear model (GLM), gradient boosting machine (GBM), random forest (RF), and eXtreme Gradient Boosting (XGBoost). All models were trained using 10-fold cross-validation on the training dataset. Model performance was evaluated based on the following metrics: area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, F-beta score, and area under the precision-recall curve (AUPRC) metrics. For comparisons between models, ANOVA was applied to normally distributed performance data, while the Kruskal–Wallis test was used for non-normally distributed variables. To further validate model generalizability, retraining was conducted on the validation set. Model interpretability and performance were assessed using the DALEX package, which generated explanatory plots and diagnostic measures. Receiver operating characteristic (ROC) curves were constructed to assess discriminatory ability. In addition, residual boxplots were plotted to visualize residual distributions, while PR curves were employed to evaluate the trade-off between precision and recall across the models.

Finally, the best-performing model was selected based on the AUC as the primary evaluation metric, supplemented by additional performance indicators. To enhance model interpretability, SHAP analysis was subsequently employed to quantify the contribution of each feature within the optimal model.

Results

Baseline characteristics

A total of 2,007 participants were included in the final analysis, of whom 1,004 were diagnosed with MAFLD. Compared to participants without MAFLD, those with MAFLD were significantly older, had a higher proportion of males, and were more likely to be of Mexican American or non-Hispanic White ethnicity. In terms of marital status, the majority of MAFLD participants were married. With respect to obesity-related indices, the MAFLD group exhibited significantly higher levels of BMI, SAT, VAT, VSR, and TAFE. Furthermore, the prevalence of hypertension and diabetes mellitus was substantially higher among participants with MAFLD compared to those without the condition (Table 1).

Development and validation of predictive models

The distributions of all candidate variables were visualized (Figures 2, 3), and inter-variable correlations were examined using Pearson correlation coefficients (Figure 4). Multicollinearity was assessed by calculating the VIF; variables exhibiting high multicollinearity—TAFE—were excluded from further analysis. Subsequently, the remaining variables were subjected to feature selection using the Boruta algorithm. This method identified the top 10 features most strongly associated with MAFLD: VAT, BMI, SAT, VSR, hypertension, diabetes mellitus, age, gender, PIR, and marital status. These features were retained for subsequent ML model development (Figure 5).

Model training was conducted in two stages. First, 10-fold cross-validation was applied to the training set to evaluate internal model

performance. Subsequently, the validation set was used to assess external performance and identify the optimal model through comparative analysis. Table 2 and Figure 6 summarized the prediction performance of the six ML models in the training set—DT, SVM, GLM, GBM, RF and XGBoost. Key evaluation metrics included the AUC, sensitivity, specificity, accuracy, F-beta score, and AUPRC, all reported as mean values. Among the six models, the GBM algorithm demonstrated the highest discriminative power in the training set, achieving the highest AUC (0.875), which was significantly superior to the other models ($p = 0.005$). It also achieved the best AUPRC (0.857, $p < 0.001$), while maintaining a favorable balance between sensitivity (0.826) and specificity (0.741). In terms of accuracy (0.784), GBM ranked jointly second with RF, following XGBoost. In the validation set (Table 3; Figure 7), XGBoost, GLM, and GBM exhibited comparable generalization performance. XGBoost achieved the highest AUC (0.882) and specificity (0.910), although its sensitivity was moderate (0.703), and it required a considerably lower optimal decision threshold (0.378) compared to GBM. Both GBM and GLM achieved AUC values of 0.879, tying for second place. GLM demonstrated the highest specificity (0.890) but the lowest sensitivity (0.717), while GBM maintained the most balanced performance with a sensitivity of 0.787 and specificity of 0.837. Residual analysis based on absolute error values (Figure 8) revealed that the GBM model exhibited a relatively narrow residual distribution, indicating greater stability and lower variance. Its median residual value was the lowest among all models, reflecting smaller average prediction errors and higher consistency. In contrast, the residuals of the XGBoost model showed greater variability, as indicated by a wider boxplot. We further evaluated the models using recall curves (Figure 9). We found that the recall curves of the GBM, XGB and RF models perform more smoothly. At high recall, GBM and XGB are able to maintain a high precision rate, while RF performs relatively stable but slightly inferior to the first two. Overall, the GBM model demonstrated superior robustness and stability. Its AUC remained consistent between the training (0.875) and validation (0.879) sets, and its sensitivity (0.826 vs. 0.787) and specificity (0.741 vs. 0.837) showed minimal fluctuation, indicating reliable and generalizable predictive performance.

SHAP interpretation of the optimal machine learning model

SHAP analysis was employed to interpret the contributions of individual features to MAFLD prediction within the optimal machine learning model (Figure 10). The SHAP summary plot revealed that VAT, BMI, and SAT were the most influential predictors, with mean absolute SHAP values of 0.187, 0.120, and 0.058, respectively. Although the VSR demonstrated a lower SHAP value (0.036), it remained among the top 10 most important features. To further elucidate the relationships between individual features and model output, SHAP dependence plots were constructed (Figure 11). These plots showed that increasing VAT and BMI values were associated with rising SHAP values, indicating a higher predicted probability of MAFLD. Among them, VAT exerted the strongest marginal effect. Although SAT contributed less than VAT and BMI, it was still a meaningful predictor. In contrast, the impact of VSR on model predictions was relatively modest, suggesting a more limited role in classification performance.

TABLE 1 Baseline population table.

Variable	Non MAFLD	MAFLD	<i>p</i> -value
<i>N</i>	1,003	1,004	
Age (years)	37.368 ± 11.557	42.744 ± 11.034	<0.001
Gender (%)			0.017
Male	479 (47.757%)	533 (53.088%)	
Female	524 (52.243%)	471 (46.912%)	
Race (%)			0.002
Mexican American	112 (11.167%)	171 (17.032%)	
Other Hispanic	87 (8.674%)	80 (7.968%)	
Non-Hispanic White	319 (31.805%)	328 (32.669%)	
Non-Hispanic Black	233 (23.230%)	195 (19.422%)	
Other Race -Including Multi-Racial	252 (25.125%)	230 (22.908%)	
Education (%)			0.021
Less than 9th grade	39 (3.888%)	57 (5.677%)	
9–11th grade	112 (11.167%)	105 (10.458%)	
High school graduate/GED or equivalent	231 (23.031%)	222 (22.112%)	
Some college or AA degree	329 (32.802%)	377 (37.550%)	
College graduate or above	292 (29.113%)	243 (24.203%)	
Marital status (%)			<0.001
Married	446 (44.467%)	534 (53.187%)	
Widowed	12 (1.196%)	18 (1.793%)	
Divorced	73 (7.278%)	100 (9.960%)	
Separated	28 (2.792%)	46 (4.582%)	
Never married	312 (31.107%)	194 (19.323%)	
Living with partner	132 (13.161%)	112 (11.155%)	
PIR	2.586 ± 1.658	2.595 ± 1.636	0.774
BMI(kg/m ²)	25.335 ± 4.891	32.546 ± 6.508	<0.001
SAT(g)	1237.871 ± 657.954	2032.317 ± 783.027	<0.001
TAFA(g)	1573.288 ± 773.010	2671.213 ± 887.613	<0.001
VAT(g)	335.418 ± 184.246	638.896 ± 249.973	<0.001
VSR	0.324 ± 0.196	0.348 ± 0.157	<0.001
Hypertension (%)			<0.001
Non Hypertension	846 (84.347%)	581 (57.869%)	
Hypertension	157 (15.653%)	423 (42.131%)	
Diabetes (%)			<0.001
Non diabetes	959 (95.613%)	827 (82.371%)	
Diabetes	44 (4.387%)	177 (17.629%)	
Smoking status (%)			0.240
Non-smoking	631 (62.911%)	606 (60.359%)	
Smoking	372 (37.089%)	398 (39.641%)	
Alcohol drinking (%)			0.933
Non-drinking	710 (70.788%)	709 (70.618%)	
Drinking	293 (29.212%)	295 (29.382%)	

PIR, Household Income-Poverty Ratio; BMI, Body mass index; VAT, Visceral adipose tissue; VAT, Visceral adipose tissue; VAT, Visceral adipose tissue; SAT, Subcutaneous adipose tissue; VSR, Visceral-to-subcutaneous fat mass ratio; TAFA, Total abdominal fat mass.

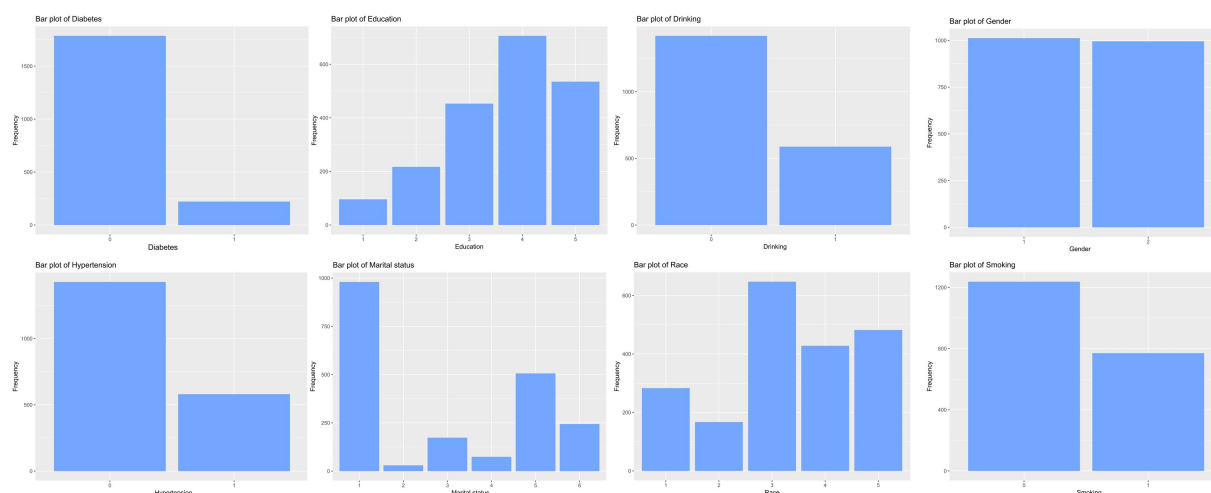


FIGURE 2
Characteristics of the distribution of categorical variables.

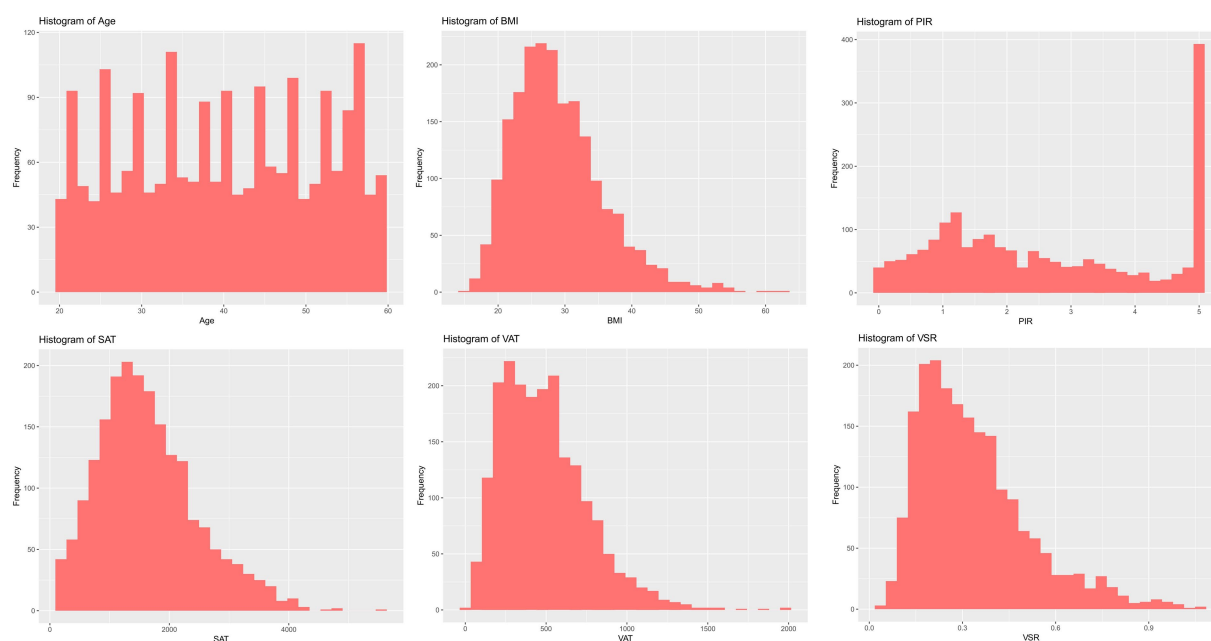


FIGURE 3
Characteristics of the distribution of continuous variables.

Discussion

In this study, we leveraged interpretable ML approaches to explore the association between body composition metrics and MAFLD using data from the 2017–2018 NHANES. Among the six ML algorithms evaluated, the GBM model demonstrated the most favorable overall performance. It achieved the highest AUC (0.879) in the validation set, closely mirroring its performance in the training set (AUC = 0.875). Additionally, the model exhibited minimal fluctuations in sensitivity and specificity across both datasets, underscoring its robustness, generalizability, and predictive reliability.

Using SHAP, we quantified the relative contributions of each selected feature to the model's predictions. VAT, BMI, and SAT emerged as the most influential predictors, highlighting the central role of abdominal fat distribution in MAFLD pathogenesis. To the best of our knowledge, this is the first study to systematically assess the predictive value of detailed body composition metrics for MAFLD using machine learning techniques. The proposed model incorporates readily obtainable demographic, lifestyle, and clinical variables, enhancing both its predictive accuracy and its potential utility in routine clinical practice and population-level screening.

Our model identified VAT, BMI, and SAT as the most important predictors of MAFLD, aligning with existing evidence on the

differential roles of adipose tissue depots in disease pathogenesis. Although BMI remains a widely used clinical measure of general obesity (19, 20), it fails to distinguish between lean mass and fat mass, and does not capture inter-individual differences in fat distribution (21). Emerging evidence suggests that the distribution of adipose tissue—particularly the accumulation of visceral fat—is more strongly associated with metabolic dysfunction and fatty liver disease than total fat mass alone (23–26). Notably, excess visceral adiposity has also been observed in individuals with normal BMI, a phenotype often referred to as “metabolically obese normal weight” or lean MAFLD. These individuals frequently exhibit greater insulin resistance and more advanced hepatic fibrosis (43), underscoring the critical role of VAT in disease progression, independent of overall body size.

Accordingly, the identification of VAT as the most important predictor in our model is biologically plausible. Visceral adipose tissue is metabolically active and, when excessively accumulated, increases the flux of FFAs into the portal circulation (44), thereby promoting

hepatic lipid deposition and inducing insulin resistance (45, 46). Additionally, VAT secretes proinflammatory cytokines such as tumor necrosis factor- α (TNF- α), interleukin-6 (IL-6), and leptin, which activate hepatic Kupffer cells and hepatic stellate cells, contributing to hepatic inflammation and fibrogenesis (47–53). In contrast, SAT functions as a metabolic buffer or “lipid reservoir.” Under conditions of energy surplus, SAT preferentially expands through adipocyte hyperplasia to safely store excess lipids and mitigate ectopic fat deposition (44, 54). However, when the storage capacity of SAT is exceeded—due to genetic, epigenetic, or adipogenic constraints—surplus lipids may overflow into visceral compartments, including the liver (24, 30, 31, 36). Therefore, the identification of SAT as an important predictive feature underscores a critical pathophysiological concept: while total adiposity contributes to metabolic burden, it is the limited expandability of SAT and the consequent visceral fat accumulation that drives the development and progression of MAFLD. Collectively, these findings provide mechanistic validation for the high predictive value of obesity-related indices in our model and illustrate the advantage of machine learning approaches in capturing the complex, interdependent relationships among metabolic risk factors.

In this study, we implemented six classical ML algorithms—DT, SVM, GLM, GBM, RF, and XGBoost—to construct predictive models for MAFLD. This multi-model approach offers a comprehensive framework for risk stratification by capturing diverse patterns of feature–outcome relationships. Each algorithm is grounded in distinct theoretical principles and exhibits unique methodological advantages. DT constructs a hierarchical decision structure via recursive binary splits, effectively modeling nonlinear feature interactions. While highly interpretable, DTs are prone to overfitting and sensitive to data noise, necessitating pruning techniques to improve generalizability. GLM, commonly applied as logistic regression, assumes linear relationships between predictors and outcomes. It provides interpretable coefficients and serves as a robust baseline model, particularly under conditions of limited sample size or when feature effects are approximately linear. SVM identifies the optimal separating hyperplane with maximum margin between classes and can incorporate nonlinear kernels to capture complex decision boundaries. It is well-suited for high-dimensional and small-sample settings,

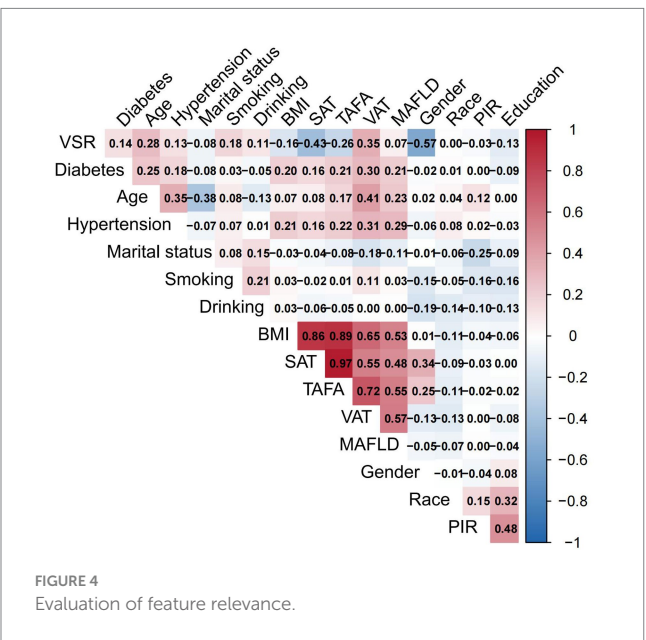


FIGURE 4
Evaluation of feature relevance.

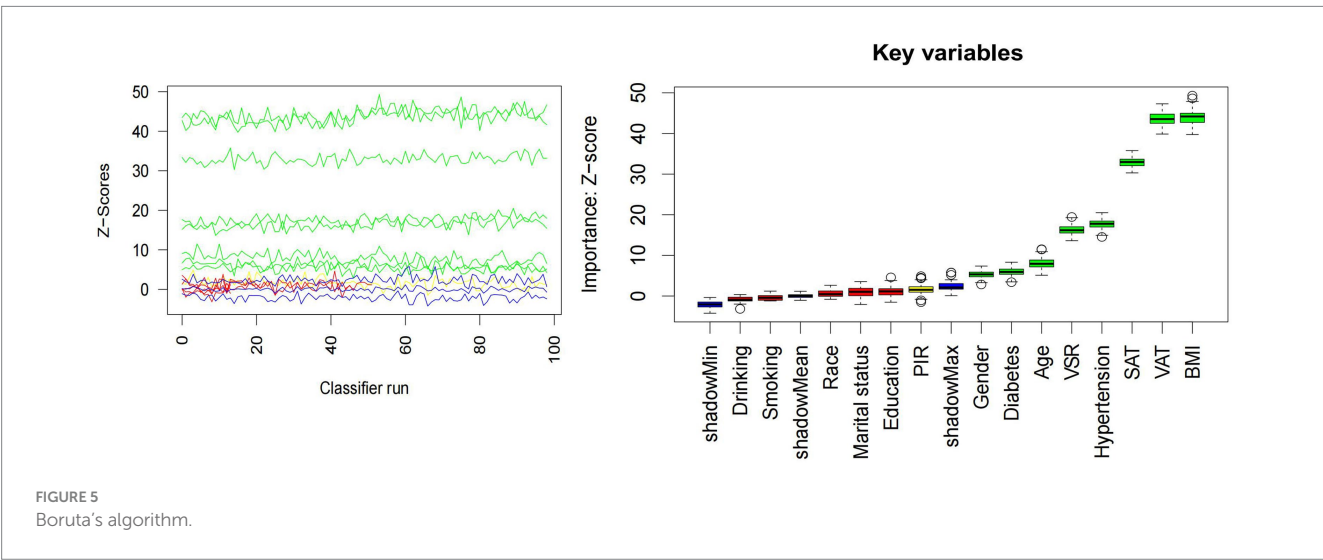


FIGURE 5
Boruta's algorithm.

TABLE 2 Six machine learning model metrics for predicting MAFLD in the training set.

Model	AUC	Specificity	Sensitivity	Accuracy	F beta	AUPRC
DT	0.803	0.697	0.859	0.778	0.795	0.750
SVM	0.867	0.752	0.804	0.778	0.784	0.850
GLM	0.871	0.785	0.758	0.772	0.769	0.850
GBM	0.875	0.741	0.826	0.784	0.792	0.857
RF	0.864	0.728	0.839	0.784	0.795	0.837
XGB	0.871	0.755	0.819	0.787	0.794	0.850
<i>p</i> -value	0.005 ^b	0.003 ^a	0.002 ^a	0.903 ^a	0.342 ^a	<0.001 ^b

DT, Decision Tree; SVM, Support Vector Machine; GLM, Generalized Linear Model; GBM, Gradient Boosting Machine; RF, Random Forest; XGB, XGBoost.

^aANOVA test.

^bKruskal-Wallis.

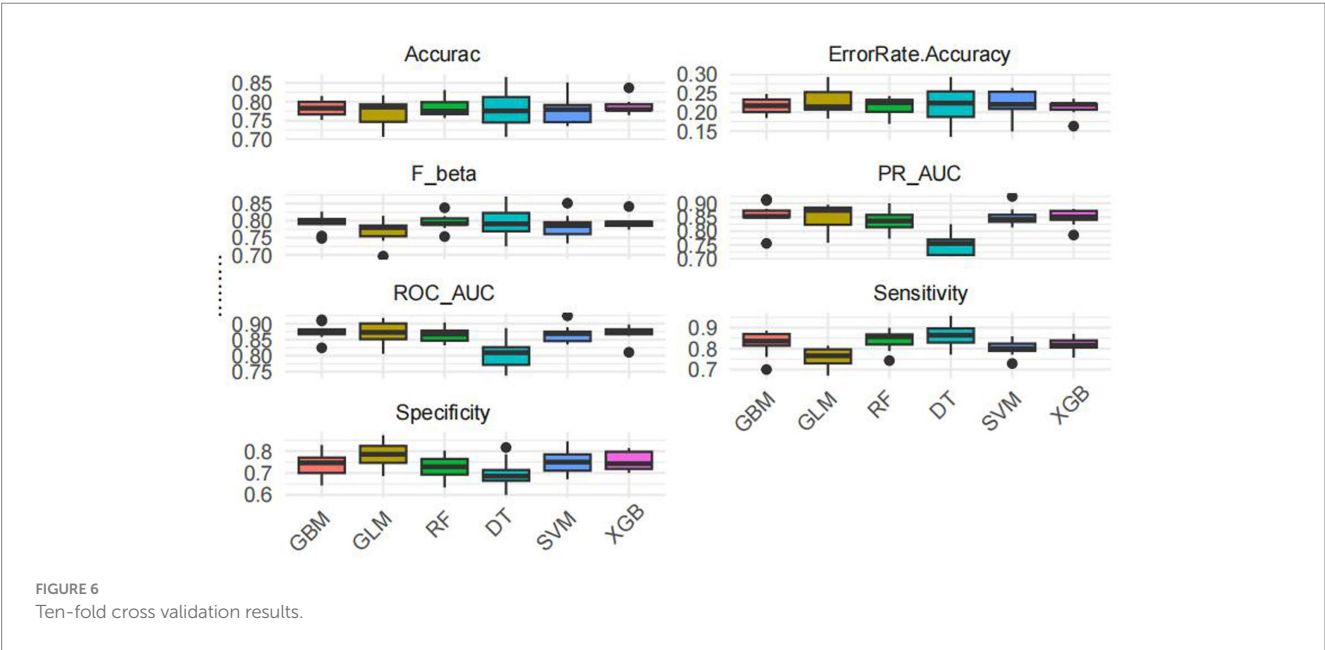


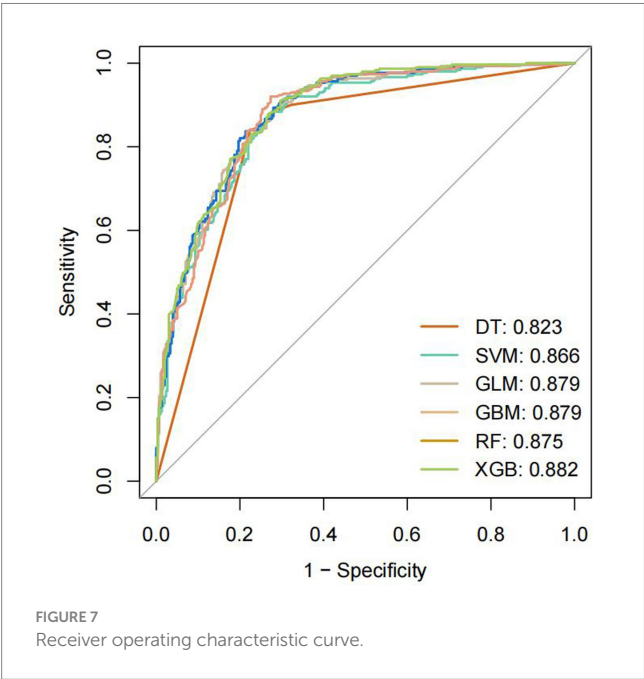
TABLE 3 Six machine learning model metrics for predicting MAFLD in the test set.

Model	AUC	Best threshold	Specificity	Sensitivity	Accuracy
DT	0.823	0.638	0.837	0.777	0.807
SVM	0.866	0.464	0.874	0.733	0.804
GLM	0.879	0.381	0.890	0.717	0.804
GBM	0.879	0.517	0.837	0.787	0.812
RF	0.875	0.413	0.920	0.727	0.824
XGB	0.882	0.378	0.910	0.703	0.807

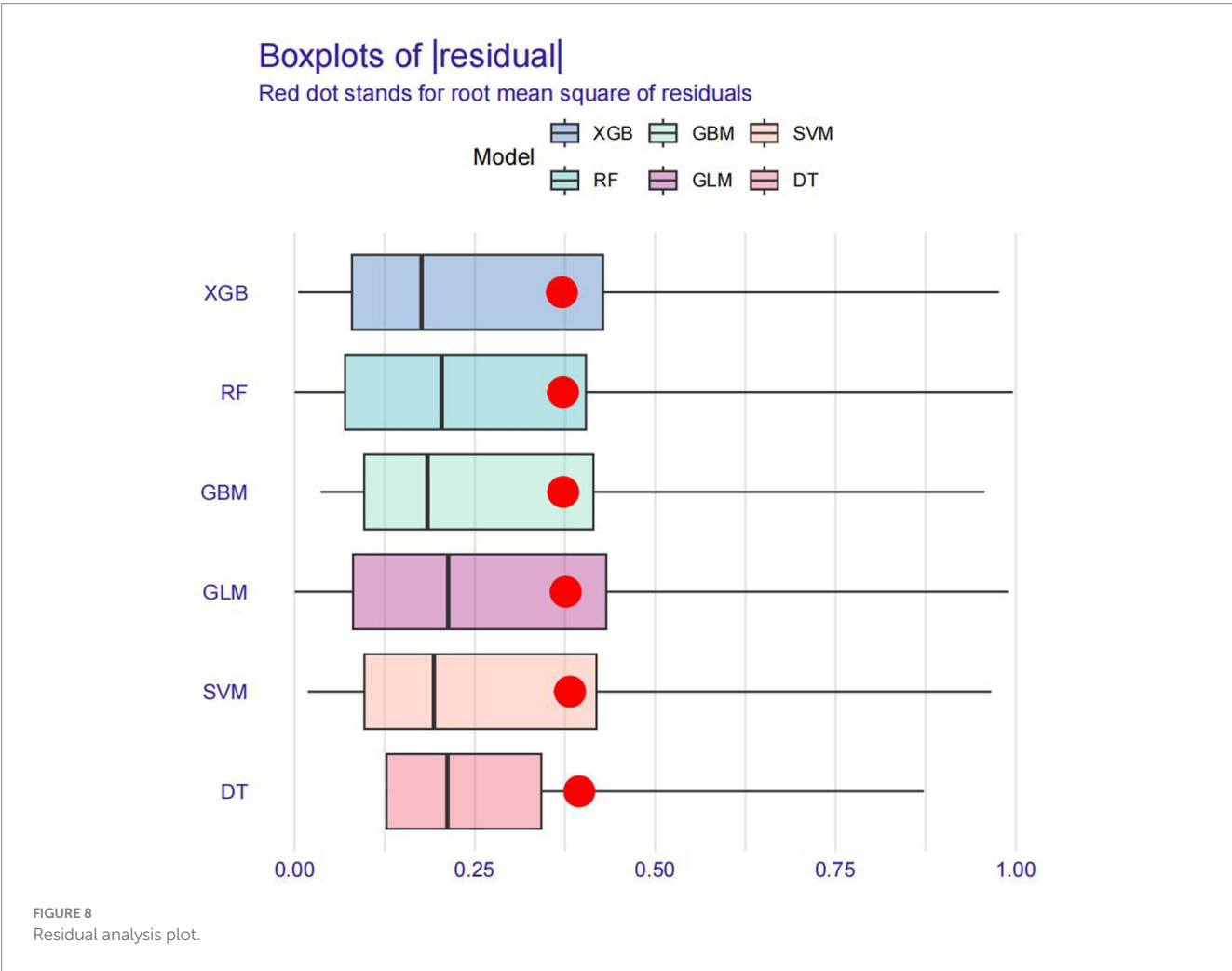
DT, Decision Tree; SVM, Support Vector Machine; GLM, Generalized Linear Model; GBM, Gradient Boosting Machine; RF, Random Forest; XGB, XGBoost.

though its model outputs are less intuitive than tree-based counterparts. RF and GBM represent ensemble learning strategies. RF leverages bagging to generate multiple decision trees trained on bootstrapped data subsets and aggregates their predictions via majority voting, thereby reducing variance and improving model stability. It also yields feature importance rankings, aiding interpretability. In contrast, GBM adopts a boosting strategy that sequentially minimizes prediction errors by fitting new models to the

residuals of prior models. This enables the modeling of intricate nonlinear relationships but requires careful hyperparameter tuning—such as tree depth and learning rate—to avoid overfitting. XGBoost, an advanced and optimized version of GBM, integrates regularization and second-order gradient approximation to enhance training efficiency, reduce overfitting, and improve predictive accuracy. It has been widely adopted across biomedical classification tasks due to its robustness and computational scalability (55).



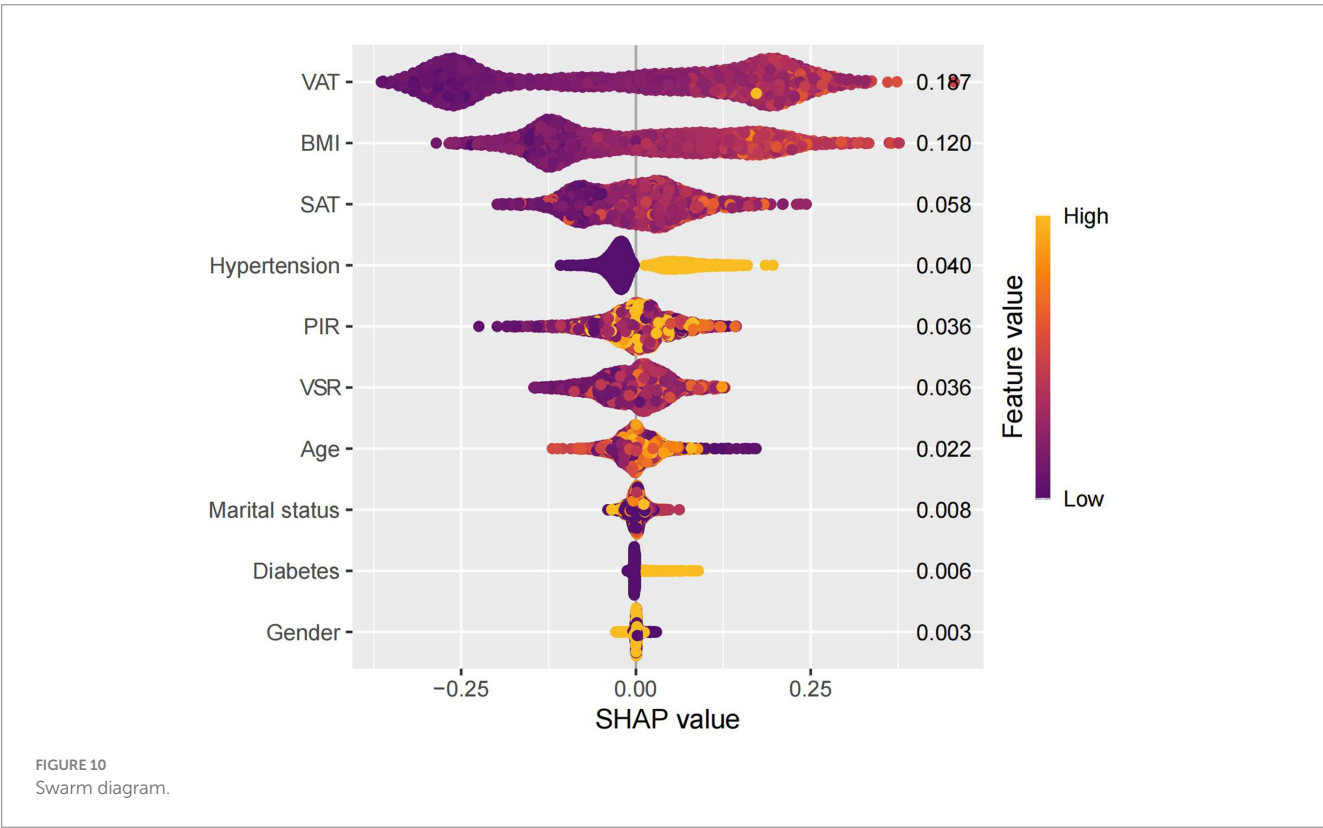
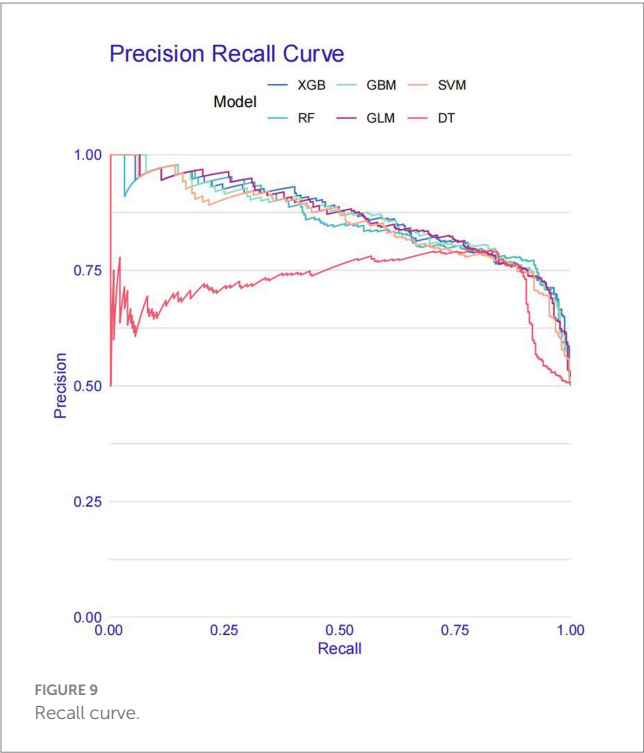
The selection of diverse machine learning algorithms for MAFLD prediction in this study is supported by both methodological rationale and prior empirical evidence. Given that MAFLD arises from a complex interplay of obesity, metabolic, and inflammation-related factors, which may exhibit nonlinear relationships and interaction effects, incorporating algorithms capable of capturing such complexities is essential. Traditional linear models may fail to identify intricate risk patterns that are better revealed by nonparametric or ensemble-based approaches. Previous studies have demonstrated the applicability and effectiveness of various ML models in fatty liver disease prediction. For instance, Qin developed decision tree, random forest, XGBoost, and support vector machine classifiers using physical examination and biochemical indicators to screen for NAFLD. Among these, the SVM model achieved the highest performance, with an AUC of approximately 0.85 and an accuracy of 80%, outperforming other models across multiple evaluation metrics (56). Similarly, Peng compared logistic regression, RF, GBM, XGBoost, and SVM models in predicting NAFLD and identified XGBoost as the top-performing algorithm, highlighting its clinical utility for early risk stratification (57). These findings support the validity of adopting multiple ML models in our framework. By leveraging the complementary strengths of different algorithms, our



approach enhances discriminative performance while maintaining robustness and interpretability—crucial attributes for translation into real-world clinical or public health applications.

This study possesses several notable strengths. First, our predictive model demonstrated satisfactory discriminative

performance in identifying individuals with MAFLD, indicating that body composition parameters can be effectively integrated into future clinical risk assessment frameworks. This finding supports the development of refined, obesity-based predictive strategies that extend beyond traditional anthropometric measures. Second, by incorporating multiple indices of body composition, our analysis underscores the critical role of fat distribution—rather than total adiposity alone—in the pathogenesis of MAFLD. This reinforces the clinical and public health imperative to shift the focus from general obesity metrics such as BMI toward a more nuanced evaluation of adipose tissue distribution. Such an approach may improve the accuracy of MAFLD screening and raise awareness of obesity-related phenotypic heterogeneity in disease development. Third, the GBM model exhibited strong translational potential. Its capacity for individualized risk estimation facilitates early identification of high-risk populations and the implementation of targeted preventive interventions aimed at mitigating progression to advanced liver disease. In resource-constrained healthcare environments, the model may assist in optimizing clinical resource allocation—for instance, by prioritizing high-risk individuals for advanced imaging modalities such as magnetic resonance imaging (MRI), thereby improving cost-effectiveness. Moreover, interpretability tools such as SHAP-derived feature importance can enhance clinician–patient communication by visualizing individual risk drivers. This may help increase patients’ understanding of their personal risk profiles and motivate adherence to lifestyle modifications, further bridging the gap between predictive analytics and actionable interventions in routine care.



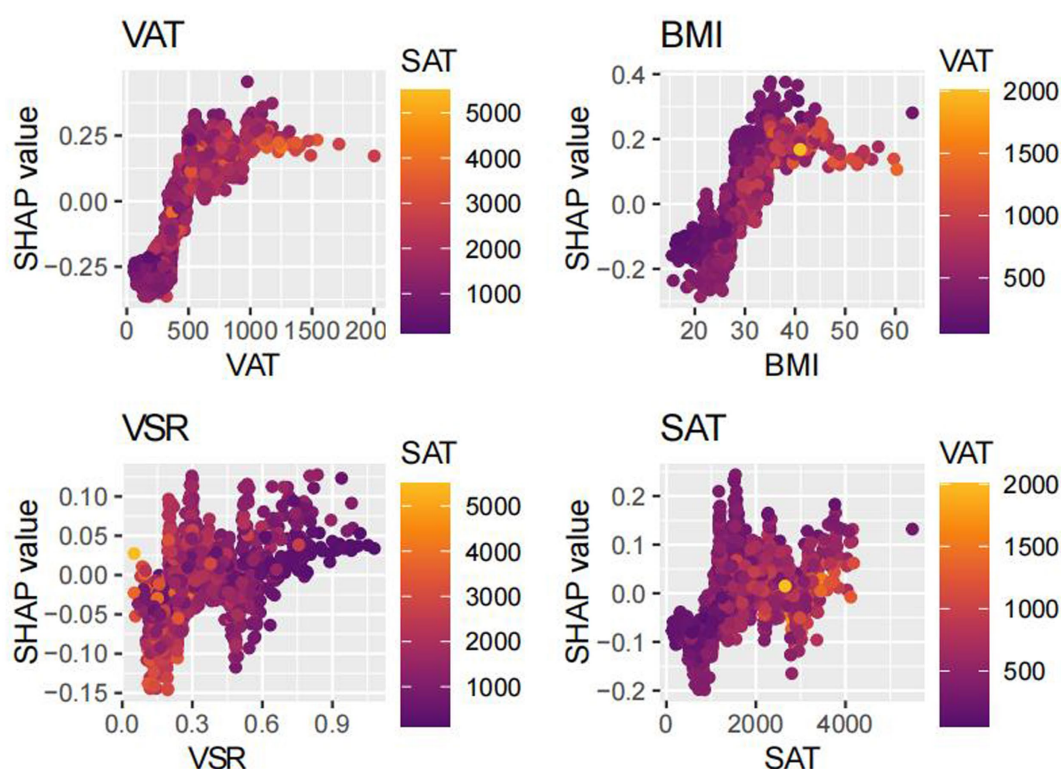


FIGURE 11
Dependency diagram.

Nevertheless, it is imperative to acknowledge the limitations inherent in the study. Firstly, histological evidence from liver biopsies is required for a definitive diagnosis of MAFLD. However, large population studies are difficult to perform invasively. Despite the utilization of FibroScan® transient elastography as an alternative in this study (a technique which has been demonstrated to possess clinical validity), the potential for the introduction of diagnostic bias remains a concern. Secondly, despite the Oral Glucose Tolerance Test (OGTT) representing a pivotal criterion for the diagnosis of diabetes, the absence of data pertaining to this indicator in the 2017–2018 NHANES cycle may have resulted in the under-recognition of cases of diabetes. Furthermore, as the data were derived from the NHANES cross-sectional survey in a single country, there are limitations in terms of sample representativeness and the applicability of the model to other populations. The model was trained and validated exclusively on an internal dataset, with no external validation performed using an independent cohort. This may be problematic due to potential differences in characteristics between populations of different races or regions, which could compromise the generalization ability of the model. Further assessment is required to ascertain the model's generalization capability. It is important to note that the cross-sectional design of this study precluded the determination of whether the observed associations were causal or not. It is therefore essential that future longitudinal studies are conducted in order to more accurately assess the causal associations between obesity indicators and the development of MAFLD. The SHAP method is predicated on the

assumption of independence of characteristics in interpreting the model. Despite the exclusion of highly correlated variables, residual correlations may still affect the interpretation of results.

Conclusion

In this study, we developed predictive models for MAFLD using six machine learning algorithms: DT, SVM, GLM, GBM, RF, and XGBoost. Among these, the GBM model demonstrated the most favorable overall performance, achieving high discriminative accuracy and stability across both training and validation datasets. Furthermore, SHAP analysis provided interpretable insights into feature contributions, with VAT emerging as the most important predictor of MAFLD risk. These findings underscore the utility of integrating advanced machine learning techniques with detailed body composition metrics to improve early risk stratification and guide targeted interventions in clinical and public health contexts.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found at: <https://wwwn.cdc.gov/nchs/nhanes/default.aspx>.

Ethics statement

The studies involving humans received ethical approval from the National Center for Health Statistics as part of the NHANES project. These studies were carried out in compliance with local laws and institutional guidelines. The participants provided their written informed consent to participate in this study.

Author contributions

YH: Formal analysis, Writing – original draft, Data curation, Methodology. XC: Writing – original draft, Data curation, Formal analysis. LW: Writing – original draft. FZ: Writing – original draft. ZZ: Writing – original draft. WX: Writing – review & editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. We acknowledge financial support from the Foshan Traditional Chinese Medicine Immune Health Technology Innovation Base, GuangDong Basic and Applied Basic Research Foundation (2023A1515140125), Foshan Joint Key Laboratory for the Research and Development and Industrialization of Chinese Medicinal Formulations, GuangDong Basic and Applied Basic Research Foundation-Natural Science Foundation

References

- Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology*. (2023) 77:1797–835. doi: 10.1097/HEP.0000000000000323
- Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology*. (2023) 78:1966–86. doi: 10.1097/HEP.0000000000000520
- Eslam M, Sarin SK, Wong VW, Wong VW-S, Fan J-G, Kawaguchi T, et al. The Asian Pacific Association for the Study of the liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. *Hepatol Int*. (2020) 14:889–919. doi: 10.1007/s12072-020-10094-2
- Mendez-Sanchez N, Arrese M, Gadano A, Oliveira CP, Fassio E, Arab JP, et al. The Latin American Association for the Study of the liver (ALEH) position statement on the redefinition of fatty liver disease. *Lancet Gastroenterol Hepatol*. (2021) 6:65–72. doi: 10.1016/S2468-1253(20)30340-X
- Shiha G, Alswat K, al M, Sharara A, Örmeci N, Waked I, et al. Nomenclature and definition of metabolic-associated fatty liver disease: a consensus from the Middle East and North Africa. *Lancet Gastroenterol Hepatol*. (2021) 6:57–64. doi: 10.1016/S2468-1253(20)30213-2
- Eslam M, Sanyal AJ, George J. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. *Gastroenterology*. (2020) 158:1999–2014. doi: 10.1053/j.gastro.2019.11.312
- Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, et al. A new definition for metabolic dysfunction-associated fatty liver disease: an international expert consensus statement. *J Hepatol*. (2020) 73:202–9. doi: 10.1016/j.jhep.2020.03.039
- Lim G, Tang A, Ng CH, Chin Y, Lim W, Tan D, et al. An observational data meta-analysis on the differences in prevalence and risk factors between MAFLD vs NAFLD. *Clin Gastroenterol Hepatol*. (2023) 21:619–29. doi: 10.1016/j.cgh.2021.11.038
- Sanyal AJ. Past, present and future perspectives in nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol*. (2019) 16:377–86. doi: 10.1038/s41575-019-0144-8
- Le MH, Yeo YH, Li X, Li X, Li J, Zou B, et al. 2019 global NAFLD prevalence: a systematic review and Meta-analysis. *Clin Gastroenterol Hepatol*. (2022) 20:2809–2817.e28. doi: 10.1016/j.cgh.2021.12.002
- (2025A1515010853), Guangdong Provincial Key Laboratory of Research and Development in Traditional Chinese Medicine (KFKT25012), and Guangdong Provincial Bureau of Traditional Chinese Medicine Scientific Research Project (Research Platform Special Program) (20254013).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

21. Vega GL, Adams-Huet B, Peshock R, Willett DW, Shah B, Grundy SM. Influence of body fat content and distribution on variation in metabolic risk. *J Clin Endocrinol Metab.* (2006) 91:4459–66. doi: 10.1210/jc.2006-0814
22. Anwar A, Rana S, Pathak P. Artificial intelligence in the management of metabolic disorders: a comprehensive review. *J Endocrinol Invest.* (2025). doi: 10.1007/s40618-025-02548-x
23. Staiano AE, Gupta AK, Katzmarzyk PT. Cardiometabolic risk factors and fat distribution in children and adolescents. *J Pediatr.* (2014) 164:560–5. doi: 10.1016/j.jpeds.2013.10.064
24. Despres JP, Lemieux I. Abdominal obesity and metabolic syndrome. *Nature.* (2006) 444:881–7. doi: 10.1038/nature05488
25. Chen X, He H, Xie K, Zhang L, Cao C. Effects of various exercise types on visceral adipose tissue in individuals with overweight and obesity: a systematic review and network meta-analysis of 84 randomized controlled trials. *Obes Rev.* (2024) 25:e13666. doi: 10.1111/obr.13666
26. Vague P. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutr.* (1956) 4:20–34. doi: 10.1093/ajcn/4.1.20
27. Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120 813 adults from 16 cohort studies from the USA and Europe. *Lancet Public Health.* (2017) 2:e277–85. doi: 10.1016/S2468-2667(17)30074-9
28. Di Angelantonio E, Bhupathiraju SN, Wormser D, Gao P, Kaptoge S, de Berrington Gonzalez A, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet.* (2016) 388:776–86. doi: 10.1016/S0140-6736(16)30175-1
29. Sahakyan KR, Somers VK, Rodriguez-Escudero JP, Hodge DO, Carter RE, Sochor O, et al. Normal-weight central obesity: implications for Total and cardiovascular mortality. *Ann Intern Med.* (2015) 163:827–35. doi: 10.7326/M14-2525
30. Elguezal Rodelo RG, Porchia LM, Torres-Rasgado E, López-Bayghen E, Gonzalez-Mejia ME. Visceral and subcutaneous abdominal fat is associated with non-alcoholic fatty liver disease while augmenting metabolic syndrome's effect on non-alcoholic fatty liver disease: a cross-sectional study of NHANES 2017–2018. *PLoS One.* (2024) 19:e298662. doi: 10.1371/journal.pone.0298662
31. Bays HE, González-Campoy JM, Henry RR, Bergman DA, Kitabchi AE, Schorr AB, et al. Is adiposopathy (sick fat) an endocrine disease? *Int J Clin Pract.* (2008) 62:1474–83. doi: 10.1111/j.1742-1241.2008.01848.x
32. Kataoka H, Nitta K, Hoshino J. Visceral fat and attribute-based medicine in chronic kidney disease. *Front Endocrinol.* (2023) 14:1097596. doi: 10.3389/fendo.2023.1097596
33. Ahima RS, Flier JS. Adipose tissue as an endocrine organ. *Trends Endocrinol Metab.* (2000) 11:327–32. doi: 10.1016/s1043-2760(00)00301-5
34. Kamada Y, Takehara T, Hayashi N. Adipocytokines and liver disease. *J Gastroenterol.* (2008) 43:811–22. doi: 10.1007/s00535-008-2213-6
35. Després JP, Lemieux I, Bergeron J, Pibarot P, Mathieu P, Larose E, et al. Abdominal obesity and the metabolic syndrome: contribution to global cardiometabolic risk. *Arterioscler Thromb Vasc Biol.* (2008) 28:1039–49. doi: 10.1161/ATVBAHA.107.159228
36. Jensen MD. Role of body fat distribution and the metabolic complications of obesity. *J Clin Endocrinol Metab.* (2008) 93:S57–63. doi: 10.1210/jc.2008-1585
37. Lind L, Strand R, Kullberg J, Ahlström H. Cardiovascular-related proteins and the abdominal visceral to subcutaneous adipose tissue ratio. *Nutr Metab Cardiovasc Dis.* (2021) 31:532–9. doi: 10.1016/j.numecd.2020.09.010
38. Deo RC. Machine learning in medicine. *Circulation.* (2015) 132:1920–30. doi: 10.1161/CIRCULATIONAHA.115.001593
39. Li W, Huang G, Tang N, Lu P, Jiang L, Lv J, et al. Effects of heavy metal exposure on hypertension: a machine learning modeling approach. *Chemosphere.* (2023) 337:139435. doi: 10.1016/j.chemosphere.2023.139435
40. Deng J, Ji W, Liu H, Li L, Wang Z, Hu Y, et al. Development and validation of a machine learning-based framework for assessing metabolic-associated fatty liver disease risk. *BMC Public Health.* (2024) 24:2545. doi: 10.1186/s12889-024-19882-z
41. Drozd K, Nabrdalik K, Kwendacz H, Hendel M, Olejarski A, Tomasik A, et al. Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: a machine learning approach. *Cardiovasc Diabetol.* (2022) 21:240. doi: 10.1186/s12933-022-01672-9
42. Eddowes PJ, Sasso M, Allison M, Tsochatzis E, Anstee QM, Sheridan D, et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology.* (2019) 156:1717–30. doi: 10.1053/j.gastro.2019.01.042
43. Bansal S, Vachher M, Arora T, Kumar B, Burman A. Visceral fat: a key mediator of NAFLD development and progression. *Hum Nutr Metab.* (2023) 33:200210. doi: 10.1016/j.hnm.2023.200210
44. Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. *Obes Rev.* (2010) 11:11–8. doi: 10.1111/j.1467-789X.2009.00623.x
45. Boden G. Obesity, insulin resistance and free fatty acids. *Curr Opin Endocrinol Diabetes Obes.* (2011) 18:139–43. doi: 10.1097/MED.0b013e3283444b09
46. Boden G, She P, Mozzoli M, Cheung P, Gumireddy K, Reddy P, et al. Free fatty acids produce insulin resistance and activate the proinflammatory nuclear factor-kappaB pathway in rat liver. *Diabetes.* (2005) 54:3458–65. doi: 10.2337/diabetes.54.12.3458
47. Mathieu P, Poirier P, Pibarot P, Lemieux I, Després JP. Visceral obesity: the link among inflammation, hypertension, and cardiovascular disease. *Hypertension.* (2009) 53:577–84. doi: 10.1161/HYPERTENSIONAHA.108.110320
48. Stepan CM, Bailey ST, Bhat S, Brown EJ, Banerjee RR, Wright CM, et al. The hormone resistin links obesity to diabetes. *Nature.* (2001) 409:307–12. doi: 10.1038/35053000
49. Vesković M, Šutulović N, Hrncić D, Stanojlović O, Macut D, Mladenović D. The interconnection between hepatic insulin resistance and metabolic dysfunction-associated Steatotic liver disease-the transition from an Adipocentric to liver-centric approach. *Curr Issues Mol Biol.* (2023) 45:9084–102. doi: 10.3390/cimb45110570
50. Johnson AM, Olefsky JM. The origins and drivers of insulin resistance. *Cell.* (2013) 152:673–84. doi: 10.1016/j.cell.2013.01.041
51. Luo JH, Wang FX, Zhao JW, Yang CL, Rong SJ, Lu WY, et al. PDIA3 defines a novel subset of adipose macrophages to exacerbate the development of obesity and metabolic disorders. *Cell Metab.* (2024) 36:2262–2280.e5. doi: 10.1016/j.cmet.2024.08.009
52. Maina V, Sutti S, Locatelli I, Vidali M, Mombello C, Bozzola C, et al. Bias in macrophage activation pattern influences non-alcoholic steatohepatitis (NASH) in mice. *Clin Sci (Lond).* (2012) 122:545–54. doi: 10.1042/CS20110366
53. Diehl KL, Vöhrer J, Hofmann K, Meiser P, Unterwiesing I, Kuerschner L, et al. Kupffer cells sense free fatty acids and regulate hepatic lipid metabolism in high-fat diet and inflammation. *Cells.* (2020) 9:2258. doi: 10.3390/cells9102258
54. Freedland ES. Role of a critical visceral adipose tissue threshold (CVATT) in metabolic syndrome: implications for controlling dietary carbohydrates: a review. *Nutr Metab.* (2004) 1:12. doi: 10.1186/1743-7075-1-12
55. Wu WT, Li YJ, Feng AZ, Li L, Huang T, Xu AD, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Mil Med Res.* (2021) 8:44. doi: 10.1186/s40779-021-00338-z
56. Qin S, Hou X, Wen Y, Wang C, Tan X, Tian H, et al. Machine learning classifiers for screening nonalcoholic fatty liver disease in general adults. *Sci Rep.* (2023) 13:3638. doi: 10.1038/s41598-023-30750-5
57. Peng HY, Duan SJ, Pan L, Wang M-Y, Chen J-L, Wang Y-C, et al. Development and validation of machine learning models for nonalcoholic fatty liver disease. *Hepatobiliary Pancreat Dis Int.* (2023) 22:615–21. doi: 10.1016/j.hbpd.2023.03.009