Journal of Advanced Research 82 (2026) 701–711

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Machine learning ensemble meets clinical practice: developing a real-world risk prediction model for metabolic syndrome using super learner and scorecard approaches

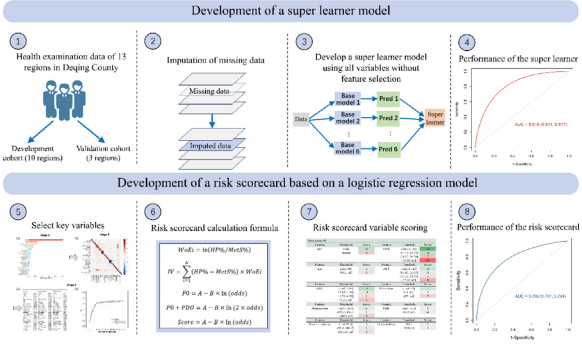
Shuwen Li[a], Yu Zhang[a], Kang Fu[b], Kailu Fang[a], Luyan Zheng[a], Yushi Lin[c], Jie Wu[a,*]

[a] *State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Centre for Infectious Diseases, Collaborative Innovation Centre for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310003, China*
[b] *Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong 258118, China*
[c] *Department of Infectious Diseases, the Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China*

HIGHLIGHTS

- Super learner model accurately predicts MetS risk using large-scale health data.
- Developed a risk scorecard for multilevel risk stratification.
- Offers a practical tool for personalized MetS risk assessment and early prevention.
- Supports dynamic risk monitoring and targeted lifestyle interventions.
- MetS score tracking motivates individuals and reinforces positive health changes.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 April 2025
Revised 18 June 2025
Accepted 24 June 2025
Available online 27 June 2025

Keywords:
Metabolic syndrome
Metabolism
Risk prediction model
Risk scores
Machine learning

ABSTRACT

Introduction: Metabolic syndrome (MetS) refers to a cluster of metabolic abnormalities that significantly increase the risk of developing cardiovascular diseases (CVDs). Traditional binary definitions of MetS fall short in capturing its severity spectrum, limiting personalized risk stratification.

Objectives: We aimed to develop a super learner model and multilevel risk scorecard to improve MetS risk prediction and support early cardiovascular risk identification.

Methods: This study included a total of 460,256 health examination records from Zhejiang, China (2018–2023), with 344,925 used for model development and 115,331 used for external validation. A super learner model combining multiple machine learning algorithms was constructed to predict MetS risk. Key predictors identified through feature selection were used to develop a logistic regression-based MetS risk scorecard, stratifying individuals into five risk levels for practical and interpretable applications.

Results: The super learner model achieved area under the receiver operating characteristics (AUCs) of 0.816 (95 % confidence interval [CI]: 0.814–0.817) and 0.810 (95 % CI: 0.808–0.813) in the development and external validation cohorts, respectively. The risk scorecard, which incorporates ten predictors, demonstrated comparable performance, with AUCs of 0.793 (95 % CI: 0.791–0.794) and 0.788 (95 % CI:

Abbreviations: AI, artificial intelligence; ALT, alanine aminotransferase; ANOVA, analysis of variance; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic; AUC, area under the curve; BMI, body mass index; BPHS, Basic Public Health Services; CHCs, community health centers; CI, confidence interval; CVDs, cardiovascular diseases; DASH, dietary approaches to stop hypertension; EMRs, electronic medical records; FBG, fasting blood glucose; HDL-C, high-density lipoprotein cholesterol; IFS, incremental feature selection; KS, Kolmogorov–Smirnov; MASLD, metabolic dysfunction-associated steatotic liver disease (nonalcoholic fatty liver disease); MetS, metabolic syndrome; PCC, Pearson correlation coefficients; SHAP, Shapley Additive exPlanation; WoE, weight of evidence.

* Corresponding author.

E-mail address: zjwujie@zju.edu.cn (J. Wu).

https://doi.org/10.1016/j.jare.2025.06.072

0.785–0.791) in the respective cohorts. Based on the risk scorecard, individuals were stratified into five categories: very low, low, normal, high, and very high risk levels.

Conclusion: The super learner model provides a highly accurate tool for MetS risk prediction, whereas the risk scorecard offers a practical and interpretable solution for clinical and personal use. These models enable precise risk assessment to guide prevention and improve outcomes.

Introduction

Metabolic syndrome (MetS) is a complex condition characterized by a cluster of metabolic abnormalities, including obesity, hyperglycemia, dyslipidemia, and hypertension [1–3]. Driven by urbanization, sedentary lifestyles, and dietary changes, the global prevalence of MetS continues to rise [1]. Additionally, MetS significantly increases the risk of cardiovascular disease and other adverse health outcomes [4]. For example, individuals with MetS are approximately twice as likely to develop cardiovascular diseases (CVDs) and 4.1 times more likely to develop type 2 diabetes than those without MetS [5]. These conditions are leading causes of disability and premature death globally and impose a significant economic burden on health care systems [6].

Diagnosis of MetS is typically based on the presence of at least three out of five key clinical indicators: increased waist circumference, increased triglycerides, decreased high-density lipoprotein cholesterol (HDL-C), increased blood pressure, and elevated fasting blood glucose (FBG) [7]. However, this binary "all-or-none" diagnostic approach overlooks the progressive nature of risk and the diverse manifestations of MetS across individuals. It is unsuitable for dynamically monitoring risk changes or assessing the effectiveness of lifestyle interventions. Given the increasing prevalence of MetS [1], developing new tools based on ensemble machine learning methods and continuous risk scores can offer a more precise and dynamic assessment of MetS risk levels.

While machine learning techniques have been widely employed for MetS prediction, most studies rely on single algorithms, which may underperform when modeling the complex relationships among risk factors or high-dimensional data [2,7,8]. To address these limitations, we developed and validated a super learner model using a large-scale health examination dataset comprising over 450,000 records from community-dwelling adults. The super learner algorithm addresses the challenge of selecting the best machine learning algorithm for a given problem, as different algorithms can yield varying results depending on the data, and the optimal choice is often unclear beforehand [9]. Given the extensive, high-dimensional data in our health examination dataset, the super learner is well suited to predict the individual risk of MetS in this context [10]. To enhance clinical operability and model interpretability, we complemented this advanced methodology with a practical risk assessment tool by developing a logistic regression-based risk scorecard that enables continuous risk scoring and multilevel risk stratification for clinical implementation. Previous studies have developed risk scorecards for large-scale screening and personal management of chronic diseases such as coronary heart disease and diabetes, demonstrating their utility in both clinical and digital health settings [11,12]. Compared with machine learning models, the scorecard offers advantages such as low computational demand, interpretability for nonspecialists, and ease of integration into electronic health records, making it suitable for use in diverse clinical and community settings. Individuals with MetS component levels near but below clinical cutoffs may not receive a diagnosis or treatment, yet still face elevated CVD risk. However, a high MetS score can help identify at-risk individuals. Prior research also indicates that individuals may be motivated by their MetS score and feel empowered when it decreases [13,14].

This study aimed to develop a robust and clinically practical MetS risk prediction model by combining a super learner ensemble algorithm for optimal predictive performance with a simplified logistic regression-based scorecard to increase interpretability and usability in real-world settings. By offering a holistic evaluation framework, these methods pave the way for more nuanced prevention strategies and targeted interventions, ultimately contributing to improved public health outcomes.

Materials and methods

Study population and design

Since 2009, China's National Basic Public Health Services Program has provided free annual health assessments to community-dwelling individuals through primary care facilities, offering standardized health examinations and follow-up consultations [15]. Leveraging this large-scale initiative, we analyzed population-based health examination data from Deqing County, Zhejiang Province, collected between January 2018 and December 2023. This county was selected for its robust implementation of the national program, ensuring high-quality, longitudinally consistent data ideal for developing predictive models. This county comprises 13 distinct community regions with a resident population exceeding 500,000. From an initial pool of 526,700 health examination records, we excluded 66,444 due to incomplete MetS component data or age under 18 years. Given that MetS components were used to define the primary outcome, missing data were not imputed to avoid introducing bias. Children and adolescents were excluded because their metabolic risk profiles are different from those of adults. The final analytical cohort included 460,256 records geographically stratified across the county's 13 regions. To ensure robust validation, we employed a random split-cohort design: 10 regions (n = 344,925) formed the derivation cohort, whereas the remaining 3 regions (n = 115,331) comprised an external validation cohort, preserving real-world generalizability.

The study was conducted in two stages. In the first stage, we developed and evaluated a comprehensive super learner model that incorporates multiple base learners and includes all available variables from the development cohort without applying feature selection. In the second stage, we identified key predictors and constructed a user-friendly MetS risk scorecard based on a logistic regression model to facilitate straightforward application in clinical practice. The study protocol conforms to the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Zhejiang University (approval number: IIT20240033B). All the collected data were anonymized and deidentified to ensure participant confidentiality. This study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guidelines **(Table S1)** [16].

Outcomes

The primary outcome was the presence of MetS, defined according to the criteria proposed in 2009 by the International Diabetes Federation and the American Heart Association/National Heart, Lung, and Blood Institute [17]. We adopted these criteria because they are widely endorsed by international guidelines and allow for ethnicity-specific thresholds, making them appropriate for Chinese adults. A diagnosis of MetS required the presence of at least three out of the following five risk factors: (1) increased waist circumference, with thresholds varying by population and country (≥90 cm for Chinese men and ≥85 cm for Chinese women) [18]; (2) elevated triglycerides (≥150 mg/dL); (3) reduced HDL-C (<40 mg/dL for men and <50 mg/dL for women); (4) elevated blood pressure (systolic blood pressure ≥130 mmHg and/or diastolic blood pressure ≥85 mmHg) or the use of antihypertensive medication; and (5) elevated FBG (≥100 mg/dL) or the use of antihyperglycemic medication.

Super learner model development and evaluation

Potential predictors were selected based on factors known to be associated with the development of MetS, drawn from both the literature and clinical expertise [8,19,20]. We included 27 variables chosen in this study to train the super learner model without additional feature selection. These variables included demographic characteristics such as age and sex; vital signs such as pulse rate and body temperature; lifestyle factors such as smoking status and alcohol consumption; and laboratory measurements such as hemoglobin and (ALT) levels. Multiple imputation by chained equations was used to handle missing values in predictor variables. Variable details can be found in the variables section of the Supplementary methods.

The super learner algorithm is illustrated in Fig. 2**a** and combines predictions from 28 distinct models included in the super learner library. These models were generated by combining six base model types—random forests, Bayesian generalized linear models, generalized linear models with elastic net regularization, multivariate adaptive regression splines, generalized additive models, and neural networks—with various parameter configurations (**Table S2**). Hyperparameter values were selected through manual tuning based on prior studies and practical performance considerations. For example, random forest models were trained with different combinations of 'mtry' (2, 5, or 8) and 'ntree' (100, 200, or 300); neural networks were tested with different node sizes (e.g., size = 5, 7, or 10) and decay parameters (e.g., 0 or 0.25); and multivariate adaptive regression splines were applied with degrees ranging from 1 to –3 (Tables S3). To train the super learner model, 10-fold cross-validation was employed within the development cohort, which means that the development cohort was divided into 10 nonoverlapping folds of approximately equal size [21]. For each fold, the base models were trained on data from the remaining nine folds, ensuring that the hold-out fold remained independent of the training process. This process was repeated across all 10 folds, allowing the generation of a complete set of predictions for each observation from all base models. The predictions from these base models were combined via optimal weights determined through cross-validation to maximize the area under the receiver operating characteristic (AUROC) curve [22]. Models with higher predictive accuracy were assigned larger weights, whereas those with lower performance received weights approaching zero. Once the optimal weights were identified, each base model was retrained on the entire development cohort to ensure robust performance. The final super learner predictions were then generated by combining these retrained models with previously determined weights [22].

We evaluated the super learner model via calibration and clinical decision curves to assess its predictive accuracy and clinical utility. Calibration curves illustrate how well the model's predicted probabilities align with observed outcomes [23]. In contrast, decision curves indicate the model's net benefit across different threshold probabilities, helping clinicians weigh the benefits of treatment decisions [24]. To improve interpretability, we used Shapley Additive exPlanation (SHAP) values, which quantify how each feature of an individual impacts their predicted risk. To further evaluate the robustness of the super learner model, we carried out subgroup analyses based on specific characteristics, including age (<60 years and ≥60 years), sex (male and female), marital status (married, never married, and divorced or widowed), and employment status (employed and unemployed).

MetS risk scorecard development and evaluation

In clinical settings, the risk scorecard offers significant advantages by providing a simplified, easy-to-use tool that facilitates decision-making and improves the usability of predictive models. By focusing on key relevant functions, the risk scorecard ensures that healthcare providers can assess risk effectively and improve the speed and accuracy of clinical decisions. Therefore, we implemented a series of feature selection methods to streamline the features included in the risk scorecard. Specifically, we employed a permutation-based importance method in which the super learner model was used to ranked the variables. This involved simulating five rounds of permutations, with each variable permuted once, and recording the changes in the area under the curve (AUC) before and after each permutation [25]. Next, we assessed the Pearson correlation coefficient (PCC) among the features; if two features had a correlation coefficient greater than 0.7, we retained the feature with higher importance, as ranked by the super learner model, and excluded the other feature. Finally, feature selection was refined via a three-step approach. First, each feature was ranked based on the least absolute shrinkage and selection operator (LASSO) regression, ridge regression, and analysis of variance (ANOVA). Second, we applied incremental feature selection (IFS), adding features one by one in order of their ranking to form various subsets. Third, each subset's predictive performance was evaluated, and the subset that achieved the best results was selected as the optimal feature set.

The scorecard was subsequently constructed using only the selected features rather than the full set used in the super learner. Continuous variables were discretized via decision tree binning, and weight of evidence (WoE) was calculated for each bin to standardize the dataset. WoE quantifies the probability of illness in each bin, providing a consistent basis for scoring [11]. The theoretical foundation of the scorecard is the logistic regression model, with the resulting scorecard resembling a financial scorecard. In line with the logic of financial scorecards, lower scores are assigned to "bad" samples, whereas in medical scorecards, lower scores are given to patients with lower risk and higher scores to those with higher risk. Therefore, in this study, individuals with MetS were categorized as "good" samples, and those without MetS were categorized as "bad" samples. The formula for calculating the risk score is defined as:

$$Score = A - B \times \ln(odds)$$

where *odds* is the ratio of good samples to bad samples and where *A* and *B* are constants.

The detailed steps and formulas for constructing the scorecard can be found in the scorecard development section of the Supplementary methods. The scorecard was calibrated to assign total scores ranging from 0 to 100, comprising a base score and bin-specific scores for each feature, reflecting the overall MetS risk for each individual. Risk stratification was conducted using the Kolmogorov–Smirnov (KS) curve to ensure effective classification of risk levels. Quintile-based categories (five groups) were selected to balance clinical interpretability with statistical discrimination, as they allow sufficient granularity to reflect risk progression while avoiding overly fragmented classification.

The risk scorecard was evaluated to ensure practical utility and clarity in clinical settings. Risk levels (very low, low, normal, high, and very high) were visualized, and pie charts were used to illustrate the distribution of characteristics within each risk level. Additionally, we examined the proportional distribution of risk levels across different demographic and clinical subgroups to identify variations in risk profiles. These evaluations provided insights into risk patterns, enabling a better understanding of MetS risk across populations.

Statistical analysis

Statistical analyses were conducted on 344,925 samples in the development cohort and 115,331 samples in the external validation cohort. Baseline characteristics are shown as the means with standard deviations for continuous variables, whereas categorical variables are reported as counts and percentages. A statistically significant difference was set at a p value of <0.05. All the statistical analyses were performed via Python (version 3.10.7) (Python Software Foundation, Wilmington, USA) and R software (version 4.3.0) (The R Foundation, Vienna, Austria).

Results

Cohort characteristics

Fig. 1 outlines the inclusion and exclusion criteria for the development and external validation cohorts. The development cohort included 344,925 individuals from 10 regions of Deqing County, with an average age of 65.80 years (SD = 9.99). Among them, 56.64 % were female, and 35.08 % met the diagnostic criteria for MetS (Table 1). For external validation, we analyzed an independent cohort of 115,331 individuals from 3 other regions. This cohort had a similar demographic profile, with an average age of 65.57 years (SD = 9.63), 57.04 % female, and a MetS incidence of 32.60 % (Table 1).

Predictive performance and clinical utility of the super learner model

The cross-validated coefficients defining the relative importance of each classifier in the super learner model are shown in Table 2. Among these, the random forest classifiers played a significant role, with coefficients of 0.172, 0.172, and 0.128 assigned to configurations of mtry = 5, ntree = 300, mtry = 8, ntree = 200, and mtry = 8, ntree = 100, respectively. Multivariate adaptive regression splines with degrees of 2 and 3 also demonstrated strong contributions, with a coefficient of 0.142. As a result, the super learner model demonstrated excellent predictive performance, achieving an AUC of 0.816 (95 % confidence interval [CI] = 0.814–0.817) in the development cohort and 0.810 (95 % CI = 0.808–0.813) in the external validation cohort **(Fig. S1)**.

To comprehensively evaluate the super learner model, we conducted a series of assessments. Using both the development and external validation cohorts, we compared the model's performance to the individual variables through metrics such as the AUC and Brier score. Results consistently demonstrated the superiority of the super learner model over single-variable predictors (Fig. 2**b**). Calibration analysis further confirmed the model's accuracy, revealing a high degree of concordance between the predicted probabilities and observed outcomes in both cohorts (Fig. 2**c**). Additionally, decision curve analysis underscored the clinical utility of the super learner model, revealing a substantial net benefit across a range of threshold probabilities (Fig. 2**c**). To enhance interpretability, we performed SHAP analysis (Fig. 2**d**). The x-axis represents the direction and magnitude of each feature's impact on the model predictions. The analysis indicated that female sex was associated with a greater risk of MetS than male sex (Fig. 2**d** and **Fig. S3a**). Furthermore, higher BMI, ALT levels, white blood cell counts, and age were strongly correlated with an increased risk of MetS (Fig. 2**d–f, Fig. S2**). Finally, subgroup analysis demonstrated the robustness of the model across various subgroups, with AUC values of approximately 0.8 **(Table S4)**.

Predictive performance and clinical utility of the risk scorecard

For feature selection in the risk scorecard, we initially employed a permutation-based importance approach using the super learner model to rank the variables (Fig. 3**a**). We subsequently calculated the PCCs between pairs of features and visualized these correlations via a heatmap (Fig. 3**b**). When the correlation coefficient between two features exceeded 0.7, we retained the more critical feature and excluded the other. For example, aspartate aminotransferase (AST) and neutrophils were excluded because of their high correlations with ALT and lymphocytes, respectively. Next, we applied three feature selection methods to rank the remaining features and utilized IFS to identify optimal subsets from the 75 (25×3) feature combinations generated (Fig. 3**c and d**). We found that ten features (BMI, sex, age, white blood cell count, alanine transaminase (ALT), pulse rate, hemoglobin, total bilirubin, comorbidities and serum creatinine) selected by both LASSO and ridge regression nearly achieved a maximum area under the curve (AUC) of 0.804. Increasing the number of features did not result in a significant improvement in AUC values. ANOVA-based selection yielded slightly lower AUCs when the feature count was the same. Notably, LASSO and ridge regression selected the same ten features, underscoring their relevance for MetS prediction.

The risk scorecard was constructed using the ten selected features. The feature scores were calculated according to the equations detailed in the scorecard development section of the Supplementary methods, with the results illustrated in Fig. 4**b**. To enhance simplicity and usability, the total score range was set to 0–100. The scorecard used a base score of 41 and variable-specific point allocations derived from binning. For example, BMI ranged from −18 ($\leq$21.6 kg/m^2) to +18 (>26.2), WBC count ranged from −8 to +5, and sex was scored as −6 (male) or +5 (female). The full details of thresholds and corresponding scores are provided in Fig. 4**b**. Individual scores were computed for each sample via the scorecard and compared with the observed outcomes to evaluate the predictive performance. In the development cohort, the scorecard yielded an AUC of 0.793 (95 % CI: 0.791–0.794), whereas in the external validation cohort, it had an AUC of 0.788 (95 % CI: 0.785–0.791) (Fig. 4**c**). Compared with the super learner model, the scorecard showed minimal performance loss, suggesting that the method used to develop the scorecard is reasonable.

To establish risk stratification, we used the KS curve to visualize the total score. A higher KS value indicates better segmentation ability of the model at the corresponding threshold. As shown in Fig. 4**a**, the maximum inflection point occurred at a score of 47, highlighting it as the optimal threshold for differentiating risk levels. For any individual tested, a lower score corresponds to a lower risk of MetS, whereas a higher score indicates a higher risk. To enhance clinical usability and standardization, the total score range was divided into five distinct risk categories: very low (0–20), low (21–40), normal (41–60), high (61–80), and very high (81–100) (Fig. 4**d**).

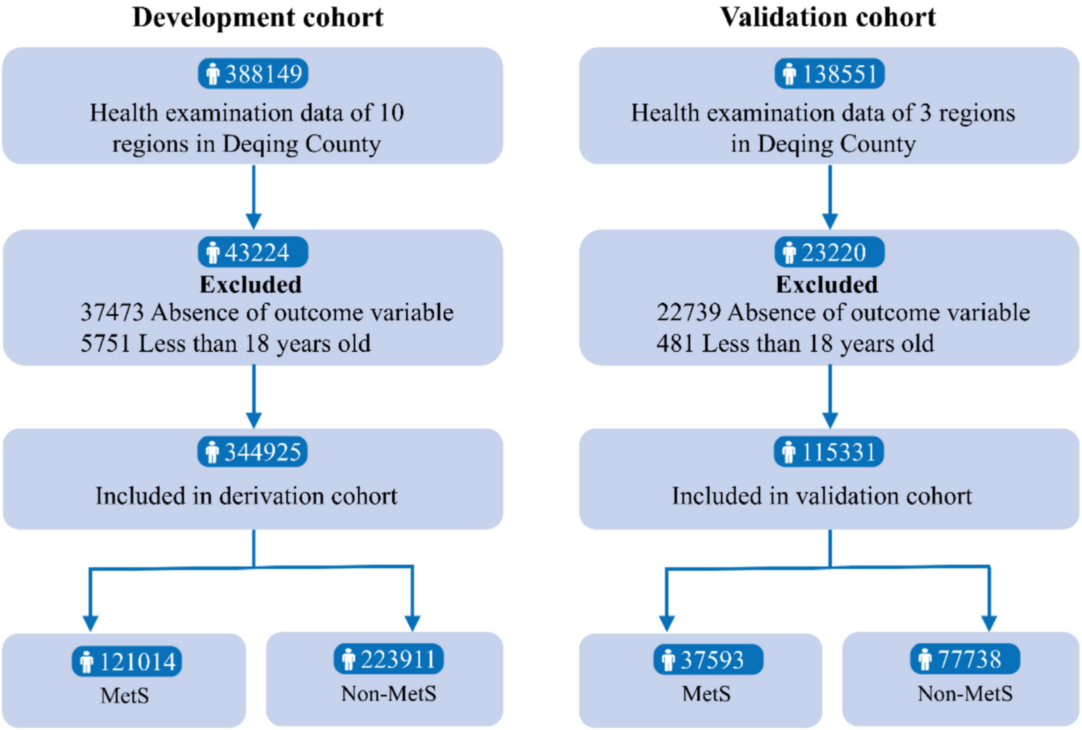


Fig. 1. Flowchart of the inclusion and exclusion criteria for the development and validation cohorts.

We further analyzed the risk distribution across different subgroups. As shown in **Fig. S4**, the risk stratification framework demonstrated consistent performance across various subgroups. Specifically, within each subgroup, the proportion of individuals diagnosed with MetS progressively declined as the risk categories shifted from “very high” to “very low”. Notably, females and elderly individuals accounted for greater proportions of the “very high” risk group. The distribution of high–risk states across demographic and clinical subgroups further underscores the utility of the scorecard (Fig. 4**e–f**). High–risk MetS states were predominantly observed among females and elderly individuals. Additionally, individuals who were divorced or widowed, unemployed, and had comorbidities presented a greater risk of MetS than those who were married, employed, and without comorbidities.

Discussion

We developed a cascade screening system for MetS risk assessment consisting of a super learner model and a risk scorecard. The system demonstrated strong predictive performance, with the super learner model achieving an AUC of 0.810, whereas the risk scorecard showed minimal performance loss and maintained relatively high predictive accuracy (AUC = 0.788). The super learner model relies solely on health examination data and can be implemented in public health strategies to help assess the prevalence of MetS in the population. The risk scorecard uses only ten simple variables and incorporates multiple risk stratifications, which may be valuable for clinical applications and personalized patient management.

MetS is a systemic disorder characterized by a cluster of metabolic abnormalities or diseases. With rapid societal development, shifting dietary patterns, sedentary lifestyles, and rising obesity rates, the prevalence of MetS has steadily increased, and younger populations are increasingly affected [26]. A significant challenge lies in raising awareness about asymptomatic metabolic abnormalities and encouraging lifestyle modifications. If left unaddressed, MetS can lead to severe long-term consequences, including a markedly increased risk of CVD, stroke, and type 2 diabetes [27,28]. In recent years, predictive models powered by big data and machine learning have shown great potential in assessing MetS risk, enabling early diagnosis, and optimizing interventions [2,7,8]. Health examination data offer clear advantages for these applications. They are easy to collect, widely available, and provide comprehensive metrics, making them ideal for large-scale disease screening. With over 450,000 records, our study represents one of the largest predictive cohorts of MetS to date. This extensive dataset provides enough statistical power to develop and validate robust models, mitigate the risk of overfitting, and enhance the ability to generalize to real-world populations undergoing routine health examinations. Given the increasing global burden of MetS and associated comorbidities [1], utilizing such datasets for risk stratification is pivotal for public health initiatives.

Our study utilizes a novel approach by integrating a super learner model with a clinically interpretable risk scorecard. The advantage of the super learner is its ability to combine predictions from multiple models (both traditional and novel), which helps avoid the potential pitfalls of relying on a single model that may perform poorly [29]. By aggregating a variety of models, ensemble methods tend to perform more robustly than single learners, approaching the “performance ceiling” for predictive accuracy [22]. In practice, however, even if a super learner outperforms a single learner, one may still prefer to use a single learner because of the computational demands of the super learner and its lack of interpretability [22]. To address this challenge, we introduced a cascading design that incorporates a super learner model followed by a logistic regression-based risk scorecard. Using the most influential features identified through a series of feature selection processes, the scorecard produced slightly lower but comparable AUC values for ease of practical use.

Traditional diagnostic methods for MetS, which rely on the presence or absence of three out of five clinical criteria, often fail to account for the progressive nature of risk and its variability

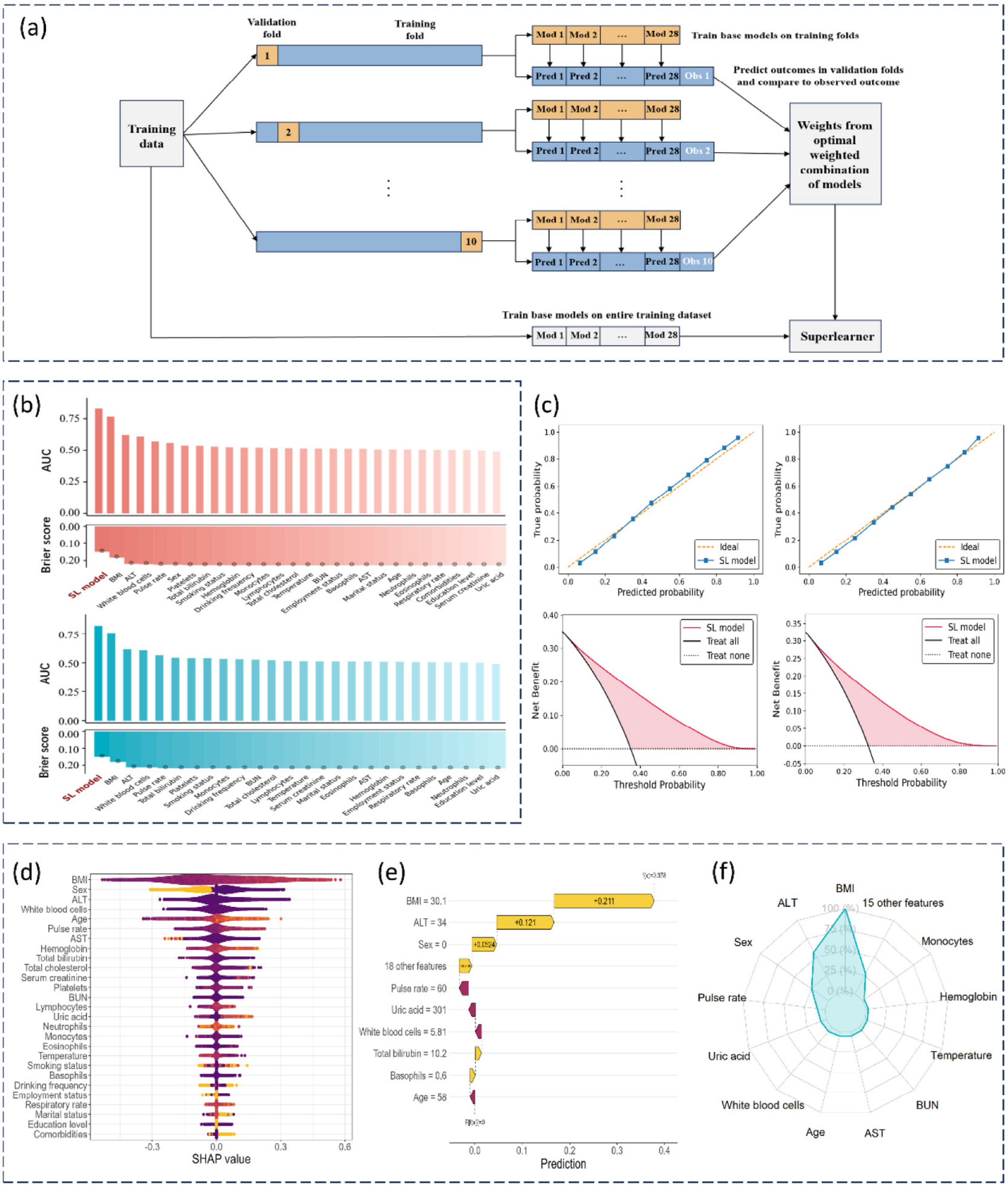


Fig. 2. Super learner model construction and model evaluation performance (a) Overview of the super learner model construction via 10-fold cross-validation with 28 base models. (b) Comparison of the predictive performance of the super learner model against 25 readily available clinical features via the AUC and Brier score in the development (n = 344,925) and external validation cohorts (n = 115,331). (c) Calibration plots and clinical decision curves for the super learner model in the development (n = 344,925) and external validation cohorts (n = 115,331). (d) SHAP summary plot showing the contribution of the predictors to the super learner model (n = 115,331). (e) Radar map illustrating the relative contribution of each predictor to MetS prediction (n = 115,331). (f) Single-sample waterfall plot quantifying the contribution of each predictor to the model output via SHAP values. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, body mass index; BUN, blood urea nitrogen; HDL-C, high-density lipoprotein cholesterol; SHAP, Shapley additive explanations; SL, super learner.

among individuals. In contrast, the risk scorecard developed in this study provides a refined stratification of individual risk, offering a more nuanced estimation of the likelihood of developing MetS. The five risk categories—very low, low, normal, high, and very high—can guide personalized management. Higher-risk individuals may benefit from closer monitoring and earlier intervention, whereas those at lower risk may require less frequent follow-up. This stratification facilitates efficient resource allocation and supports individualized preventive strategies. The scorecard is designed for both digital formats, such as integration into electronic health records, and paper-based formats for use in community health settings. As demonstrated in previous studies [13,30,31], a decreasing MetS score can serve as a motivating factor, encouraging individuals to make healthier choices. This study highlights several modifiable factors associated with MetS, including BMI, white blood cell counts, ALT, pulse rate, hemoglobin, total bilirubin, comorbidities, and serum creatinine. For instance, BMI is a key factor linked to lifestyle choices. Evidence suggests that comprehensive lifestyle interventions, including dietary changes such as the Dietary Approaches to Stop Hypertension (DASH) or Mediterranean diets, moderate caloric restriction, and regular physical activity, can significantly improve BMI and reduce the risk of MetS [32–35]. Comorbidities such as cardiovascular, cerebrovascular, and kidney disorders exacerbate inflammation and insulin resistance, disrupt-

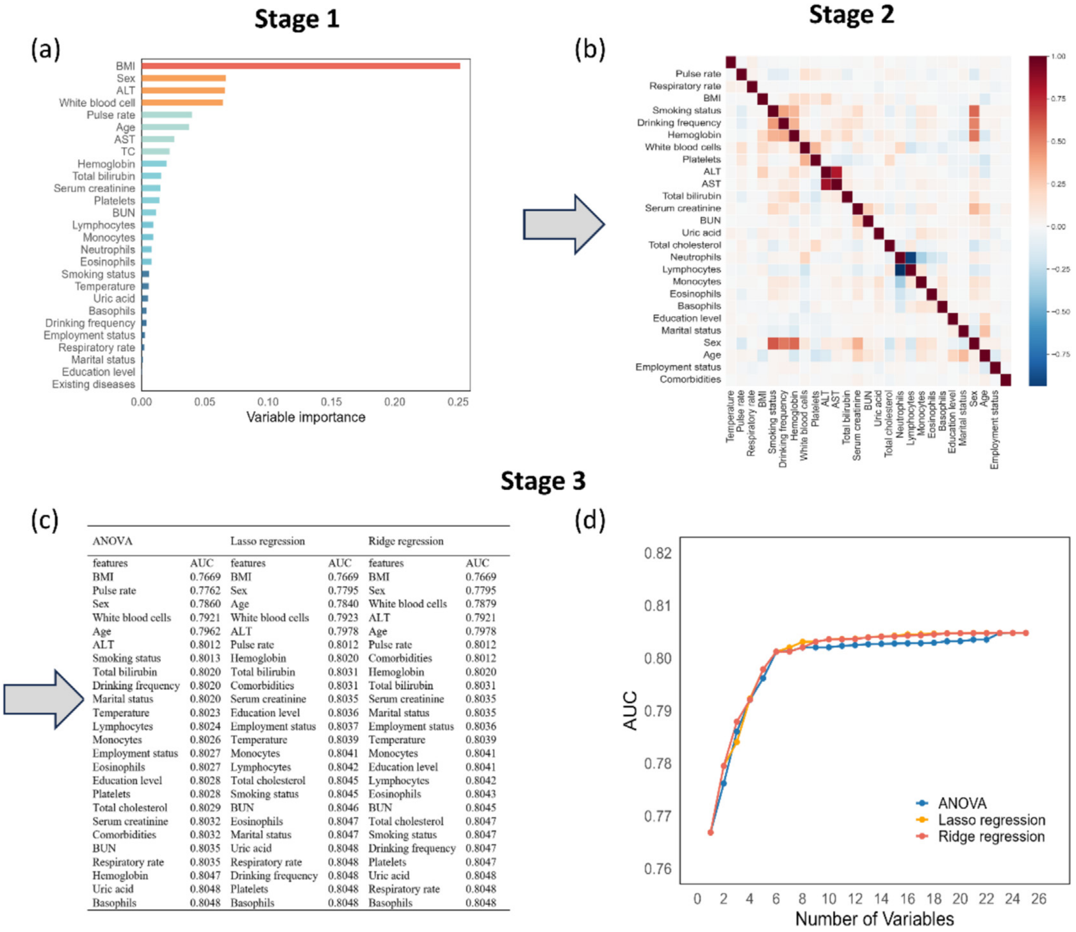
Table 1
Baseline characteristics of the development and external validation cohorts.

Variables	Development cohort(n = 344925)	External validation cohort (n = 115,331)
Demographic		
Age (y)	65.80 (9.99)	65.57 (9.63)
Sex		
Male	149,561 (43.36)	49,549 (42.96)
Female	195,364 (56.64)	65,782 (57.04)
Education level		
Less than middle school	329,617 (95.56)	112,513 (97.56)
More than middle school	15,308 (4.44)	2818 (2.44)
Employment status		
Employed	253,289 (73.43)	103,086 (89.38)
Unemployed	91,636 (26.57)	12,245 (10.62)
Marital status		
Never married	3204 (0.93)	888 (0.77)
Married	310,287 (89.96)	103,867 (90.06)
Divorced or widowed	31,434 (9.11)	10,576 (9.17)
Waist circumference (cm)	83.34 (9.49)	82.90 (9.46)
Systolic blood pressure (mmHg)	138.83 (19.06)	137.22 (18.47)
Diastolic blood pressure (mmHg)	81.97 (10.14)	82.53 (9.88)
BMI (kg/m^2)	23.82 (3.39)	23.77 (3.30)
Vital signs		
Pulse rate (per minute)	72.61 (11.70)	74.86 (10.56)
Respiratory rate (per minute)	17.96 (1.02)	17.78 (1.06)
Temperature (°C)	36.59 (0.26)	36.58 (0.25)
Health lifestyle behavior		
Smoking status		
Never smoked	265,102 (76.86)	88,980 (77.15)
Smoking	63,423 (18.39)	22,650 (19.64)
Smoking cessation	16,400 (4.75)	3701 (3.21)
Drinking frequency		
Never	265,041 (76.84)	91,217 (79.09)
Occasionally	15,569 (4.51)	4639 (4.02)
Often	5681 (1.65)	3214 (2.79)
Everyday	58,634 (17.00)	16,261 (14.10)
Laboratory analysis		
Hemoglobin (g/L)	137.41 (14.64)	134.38 (14.21)
White blood cells (10^9/L)	5.65 (1.64)	5.56 (1.52)
Platelets (10^9/L)	192.55 (57.11)	187.36 (56.83)
Fasting blood glucose (mg/dL)	100.00 (25.47)	97.65 (26.72)
ALT (U/L)	20.67 (16.05)	20.60 (16.58)
AST (U/L)	23.55 (11.42)	23.56 (12.47)
Total bilirubin (μmol/L)	14.36 (6.68)	13.74 (6.45)
Serum creatinine (μmol/L)	70.26 (27.43)	69.62 (26.65)
BUN (mmol/L)	5.77 (3.08)	5.73 (1.67)
Uric acid (μmol/L)	44.90 (115.48)	6.57 (15.27)
Total cholesterol (mmol/L)	4.94 (1.00)	4.86 (0.94)
Triglycerides (mg/dL)	141.02 (108.27)	143.33 (108.66)
HDL-C (mg/dL)	52.74 (13.33)	52.22 (12.67)
Neutrophils	58.94 (9.79)	59.23 (8.55)
Lymphocytes	32.29 (9.12)	32.02 (8.16)
Monocytes	5.83 (2.06)	5.89 (1.65)
Eosinophils	2.52 (2.52)	2.49 (1.99)
Basophils	0.42 (0.32)	0.38 (0.24)
Comorbidities		
Without comorbidities	341,400 (98.98)	114,269 (99.08)
With comorbidities	3525 (1.02)	1062 (0.92)
Cerebrovascular disease	911 (0.26)	237 (0.21)
Kidney disease	200 (0.06)	121 (0.10)
Heart disease	1081 (0.31)	451 (0.39)
Vascular disease	53 (0.02)	15 (0.01)
Eye disease	1088 (0.32)	135 (0.12)
Neurological disease	299 (0.09)	141 (0.12)
Metabolic syndrome	121,014 (35.08)	37,593 (32.60)

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; BUN, blood urea nitrogen; HDL-C, high-density lipoprotein cholesterol.

Table 2
Super learner risk and coefficients for candidate learners with nonzero coefficients.

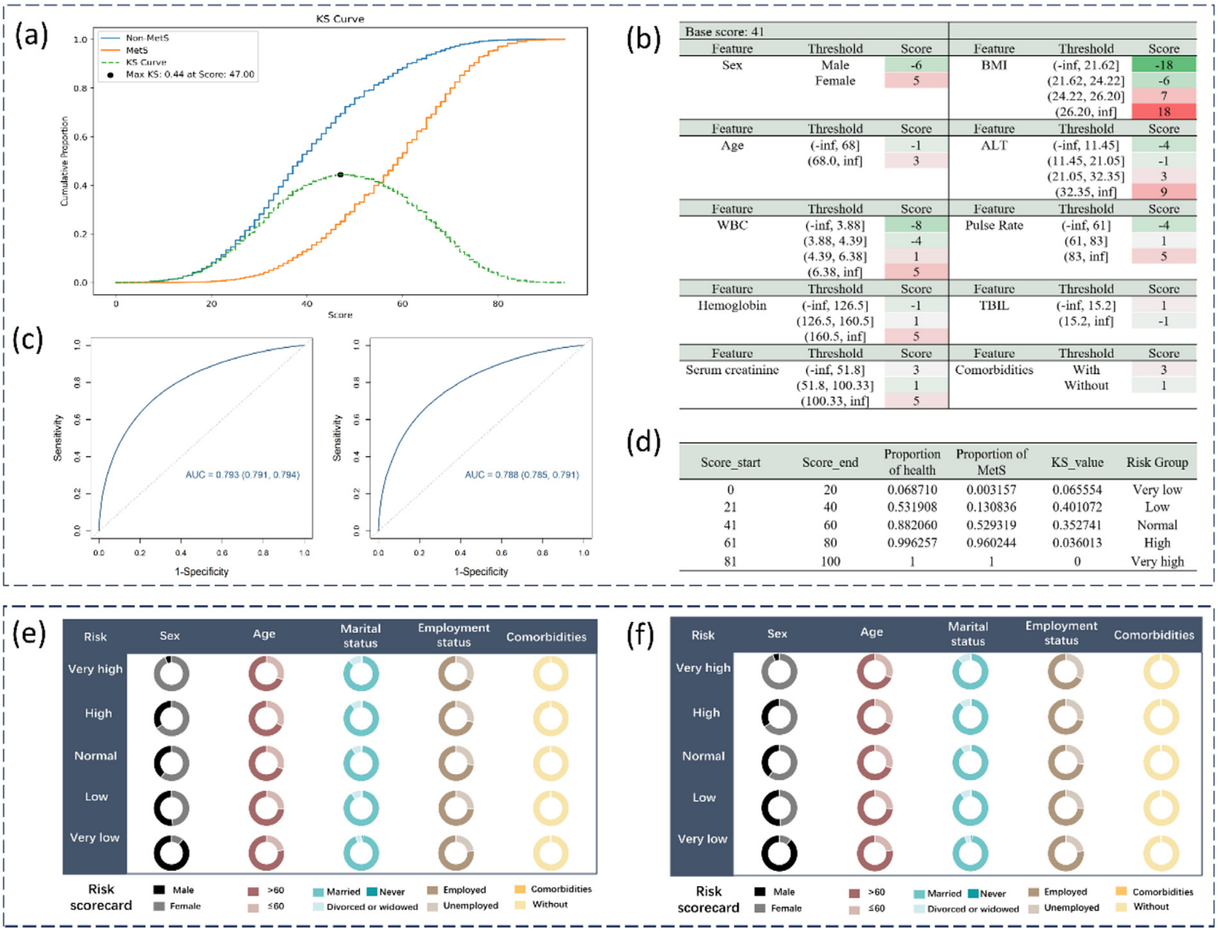
Model	Hyperparameter setting	Risk	Coefficients (β)
Multivariate adaptive regression splines	degree = 2	0.186	0.142
Multivariate adaptive regression splines	degree = 3	0.186	0.142
Generalized additive model	Default	0.189	0.006
Neural network	size = 10, decay = 0.5	0.191	0.071
Neural network	size = 5, decay = 0.25	0.421	0.010
Neural network	size = 9, decay = 0	0.301	0.011
Random forest	mtry = 2, ntree = 100	0.191	0.018
Random forest	mtry = 2, ntree = 200	0.186	0.025
Random forest	mtry = 2, ntree = 300	0.188	0.043
Random forest	mtry = 5, ntree = 100	0.185	0.082
Random forest	mtry = 5, ntree = 200	0.186	0.025
Random forest	mtry = 5, ntree = 300	0.186	0.172
Random forest	mtry = 8, ntree = 100	0.184	0.128
Random forest	mtry = 8, ntree = 200	0.186	0.172
Random forest	mtry = 8, ntree = 300	0.185	0.091



ANOVA		Lasso regression		Ridge regression	
features	AUC	features	AUC	features	AUC
BMI	0.7669	BMI	0.7669	BMI	0.7669
Pulse rate	0.7762	Sex	0.7795	Sex	0.7795
Sex	0.7860	Age	0.7840	White blood cells	0.7879
White blood cells	0.7921	White blood cells	0.7923	ALT	0.7921
Age	0.7962	ALT	0.7978	Age	0.7978
ALT	0.8012	Pulse rate	0.8012	Pulse rate	0.8012
Smoking status	0.8013	Hemoglobin	0.8020	Comorbidities	0.8012
Total bilirubin	0.8020	Total bilirubin	0.8031	Hemoglobin	0.8020
Drinking frequency	0.8020	Comorbidities	0.8031	Total bilirubin	0.8031
Marital status	0.8020	Serum creatinine	0.8035	Serum creatinine	0.8035
Temperature	0.8023	Education level	0.8036	Marital status	0.8035
Lymphocytes	0.8024	Employment status	0.8037	Employment status	0.8036
Monocytes	0.8026	Temperature	0.8039	Temperature	0.8039
Employment status	0.8027	Monocytes	0.8041	Monocytes	0.8041
Eosinophils	0.8027	Lymphocytes	0.8042	Education level	0.8041
Education level	0.8028	Total cholesterol	0.8045	Lymphocytes	0.8042
Platelets	0.8028	Smoking status	0.8045	Eosinophils	0.8043
Total cholesterol	0.8029	BUN	0.8046	BUN	0.8045
Serum creatinine	0.8032	Eosinophils	0.8047	Total cholesterol	0.8047
Comorbidities	0.8032	Marital status	0.8047	Smoking status	0.8047
BUN	0.8035	Uric acid	0.8048	Drinking frequency	0.8047
Respiratory rate	0.8035	Respiratory rate	0.8048	Platelets	0.8047
Hemoglobin	0.8047	Drinking frequency	0.8048	Uric acid	0.8048
Uric acid	0.8048	Platelets	0.8048	Respiratory rate	0.8048
Basophils	0.8048	Basophils	0.8048	Basophils	0.8048

Fig. 3. Workflow for feature selection in constructing a simple risk scorecard (a) Variable importance ranking of predictors based on the super learner model (n = 344,925). (b) Heatmap showing the Pearson correlation coefficients among features (n = 344,925). (c) Feature selection via three statistical methods in the IFS strategy (n = 344,925). (d) IFS results comparing the performance of three feature selection techniques (n = 344,925). Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, body mass index; BUN, blood urea nitrogen; HDL-C, high-density lipoprotein cholesterol.

ing metabolic homeostasis through metabolic abnormalities and endocrine dysfunction [36,37]. Early screening and timely intervention for these conditions can help prevent metabolic imbalance and slow MetS progression. Interventions aimed at reducing an individual's risk score or modifying its components are likely to decrease the overall risk of MetS. Furthermore, the risk scorecard has potential applications in tailoring therapeutic decisions based on individual risk levels and monitoring the effectiveness of these interventions. By providing actionable insights, this tool can support both preventive strategies and personalized management of MetS. In addition to individualized care, the scorecard could also support broader population health strategies. For example, it can assist community clinics in prioritizing high-risk individuals for screening or follow-up and informing triage decisions in



Base score: 41

Feature	Threshold	Score	Feature	Threshold	Score
Sex	Male	-6	BMI	(-inf, 21.62]	-18
	Female	5		(21.62, 24.22]	-6
				(24.22, 26.20]	7
				(26.20, inf]	18
Feature	Threshold	Score	Feature	Threshold	Score
Age	(-inf, 68]	-1	ALT	(-inf, 11.45]	-4
	(68.0, inf]	3		(11.45, 21.05]	-1
				(21.05, 32.35]	3
				(32.35, inf]	9
Feature	Threshold	Score	Feature	Threshold	Score
WBC	(-inf, 3.88]	-8	Pulse Rate	(-inf, 61]	-4
	(3.88, 4.39]	-4		(61, 83]	1
	(4.39, 6.38]	1		(83, inf]	5
	(6.38, inf]	5			
Feature	Threshold	Score	Feature	Threshold	Score
Hemoglobin	(-inf, 126.5]	-1	TBIL	(-inf, 15.2]	1
	(126.5, 160.5]	1		(15.2, inf]	-1
	(160.5, inf]	5			
Feature	Threshold	Score	Feature	Threshold	Score
Serum creatinine	(-inf, 51.8]	3	Comorbidities	With	3
	(51.8, 100.33]	1		Without	1
	(100.33, inf]	5			

Score_start	Score_end	Proportion of health	Proportion of MetS	KS_value	Risk Group
0	20	0.068710	0.003157	0.065554	Very low
21	40	0.531908	0.130836	0.401072	Low
41	60	0.882060	0.529319	0.352741	Normal
61	80	0.996257	0.960244	0.036013	High
81	100	1	1	0	Very high

Fig. 4. Construction of the MetS scorecard, risk stratification, and group characteristics (a) KS curve for the MetS risk scorecard (n = 344,925). (b) Visualization of the MetS risk scorecard. (c) ROC curves for the MetS risk scorecard in the development (n = 344,925) and external validation cohorts (n = 115,331). (d) Thresholds for risk stratification groups based on the MetS scorecard. (e) Pie chart showing the baseline characteristics of the risk groups in the development cohort (n = 344,925). (f) Pie chart showing the baseline characteristics of the risk groups in the external validation cohort (n = 115,331). Abbreviations: ALT, alanine aminotransferase; AUC, area under the curve; BMI, body mass index; TBIL, total bilirubin; WBC, white blood cell count.

resource-limited settings. Its compatibility with electronic health systems or integration into routine check-ups enables timely intervention without requiring specialist input. These features make it a practical tool for implementing scalable, data-driven MetS prevention and management programs.

The subgroup analysis of the super learner model demonstrated stable predictive performance across all examined subgroups in both the development and external validation cohorts, with most AUCs exceeding 0.80 and only minor variations observed in sensitivity and specificity. We also assessed how the risk scorecard stratified individuals across the same subgroups. The results revealed that those in the “very high risk” category were more likely to be older, female, unemployed, and to have comorbidities. The SHAP analysis of the super learner model also indicated that female sex was associated with an increased risk of MetS. This finding can be partly explained by the age distribution of our study population, where the mean age exceeded 65 years. A substantial proportion of the female participants were therefore likely to be postmenopausal. The decline in estrogen after menopause is known to contribute to an increase in visceral adiposity, insulin resistance, and dyslipidemia, which collectively increase the risk of MetS [38–40]. In addition, older adults tend to have metabolic changes, lower physical activity, and more chronic conditions, all of which increase MetS risk [41]. Unemployment may indicate lower socioeconomic status, which is often linked to limited healthcare access, poor diet, and psychological stress [42]. Comorbidities such as cardiovascular or renal diseases may further increase risk due to the underlying metabolic burden [36,37]. These findings support the scorecard’s ability to stratify risk across diverse real-world populations.

This study has several limitations that should be noted. First, the performance of the super learner depends on the inclusion of the optimal algorithm within its library. While our super learner incorporates a diverse range of adaptive and smooth learners, it is possible that some algorithms outside the current selection could perform better. Future research may consider expanding the library to explore additional methods. Second, the absence of dietary data and other behavioral variables may restrict the comprehensiveness of our prediction model. In ensemble models, adding diverse and informative predictors often increases AUC performance. Future versions of the model may benefit from integrating these variables to capture a more complete metabolic profile. Third, participants without complete MetS information were excluded because imputation was not possible. Nevertheless, the model was developed with a large sample size and externally validated, making it reasonable to assume that the missing data had minimal impact on the overall findings. Fourth, given the relatively advanced age of our study cohort (mean age ~65 years), caution should be exercised when generalizing the findings to younger populations. Future research should explore age-specific recalibration or conduct external validation in younger cohorts to improve generalizability. Finally, future studies are needed to further validate the model in more diverse regions, such as highly urbanized cities or rural areas in western China, where socioeconomic development and lifestyle patterns differ significantly.

In conclusion, we developed a novel cascading screening system for MetS risk evaluation that combines advanced super learner methods with a practical risk scorecard. This approach demonstrates robust predictive performance and offers significant potential for public health applications and clinical decision-making. Future studies should focus on validating the model in multiethnic and urban populations, where metabolic profiles and healthcare access may vary significantly. Additionally, integrating physiological data from wearable devices may enable real-time, dynamic risk monitoring. Further research is needed to evaluate the cost-effectiveness and clinical impact of deploying the model in real-world primary care settings. By integrating cutting-edge analytics with real-world implementation, this study provides a roadmap for enhancing MetS risk prediction and management.

Authors' contributions

J.W. designed the study and supervised the overall process. S.W. L. accessed and verified all the data. S.W.L., Y.Z., and K.F. analyzed the data and interpreted the results. S.W.L., Y.Z., K.F., K.L.F., L.Y.Z., and Y.S.L. wrote the first draft of the manuscript. J.W. reviewed and edited the manuscript. All authors contributed to discussion, reviewed, and approved the final version of the manuscript. J.W. is responsible for the decision to submit the manuscript.

Compliance with ethics requirement

The study protocol conforms to the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Zhejiang University (approval number: IIT20240033B).

Funding

This study was funded by National Natural Science Foundation of China (72374179, 71904170), the Fundamental Research Funds for the Central Universities (2025ZFJH03), Zhejiang University K. P. Chao's High Technology Development Foundation (2022RC017), Mega-Project of National Science and Technology for the 13th Five-Year Plan of China (2018ZX10721102-003-006, 2018ZX10715013-003-003), and Zhejiang Province Healthcare Innovation Talent Program.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We highly appreciate the financial support from National Natural Science Foundation of China (72374179,71904170), the Fundamental Research Funds for the Central Universities (2025ZFJH03), Zhejiang University K. P. Chao's High Technology Development Foundation (2022RC017), Mega-Project of National Science and Technology for the 13th Five-Year Plan of China (2018ZX10721102-003-006, 2018ZX10715013-003-003), and Zhejiang Province Healthcare Innovation Talent Program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2025.06.072.

References

[1] Neeland IJ et al. Metabolic syndrome. Nat Rev Dis Primers 2024;10:77. doi: https://doi.org/10.1038/s41572-024-00563-5.
[2] Kim J, Mun S, Lee S, Jeong K, Baek Y. Prediction of metabolic and pre-metabolic syndromes using machine learning models with anthropometric, lifestyle, and biochemical factors from a middle-aged population in Korea. BMC Public Health 2022;22:664. doi: https://doi.org/10.1186/s12889-022-13131-x.
[3] Zhang F et al. Association of metabolic syndrome and its components with risk of stroke recurrence and mortality: a meta-analysis. Neurology 2021;97: e695–705. doi: https://doi.org/10.1212/WNL.0000000000012415.
[4] Dai X, Feng S, Li T. Cold atmospheric plasma control metabolic syndromes via targeting fat mass and obesity-associated protein. Pharmacol Res 2025;215:107720. doi: https://doi.org/10.1016/j.phrs.2025.107720.
[5] Yang S et al. Development and validation of an age-sex-ethnicity-specific metabolic syndrome score in the Chinese adults. Nat. Commun. 2023;14. doi: https://doi.org/10.1038/s41467-023-42423-y.
[6] Mohseni-Takalloo S, Mozaffari-Khosravi H, Mohseni H, Mirzaei M, Hosseinzadeh M. Metabolic syndrome prediction using non-invasive and dietary parameters based on a support vector machine. Nutr Metab Cardiovasc Dis 2024;34:126–35. doi: https://doi.org/10.1016/j.numecd.2023.08.018.
[7] Daniel Tavares L et al. Prediction of metabolic syndrome: a machine learning approach to help primary prevention. Diabetes Res Clin Pract 2022;191:110047. doi: https://doi.org/10.1016/j.diabres.2022.110047.
[8] Baichoo S, Shin H, Shim S, Oh S. Machine learning-based predictive model for prevention of metabolic syndrome. PLoS One 2023;18. doi: https://doi.org/10.1371/journal.pone.0286635.
[9] Wardenaar KJ et al. Common and specific determinants of 9-year depression and anxiety course-trajectories: a machine-learning investigation in the Netherlands Study of Depression and anxiety (NESDA). J Affect Disord 2021;293:295–304. doi: https://doi.org/10.1016/j.jad.2021.06.029.
[10] Butzin-Dozier Z et al. Predicting long COVID in the national COVID cohort collaborative using super learner: cohort study. JMIR Public Health Surveill 2024;10:e53322. doi: https://doi.org/10.2196/53322.
[11] Yang H et al. A gender specific risk assessment of coronary heart disease based on physical examination data. NPJ Digit Med 2023;6:136. doi: https://doi.org/10.1038/s41746-023-00887-8.
[12] Yang H et al. Risk prediction of diabetes: big data mining with fusion of multifarious physical examination indicators. Inf Fusion 2021;75:140–9. doi: https://doi.org/10.1016/j.inffus.2021.02.015.
[13] DeBoer MD, Filipp SL, Gurka MJ. Use of a metabolic syndrome severity Z score to track risk during treatment of prediabetes: an analysis of the diabetes prevention program. Diabetes Care 2018;41:2421–30. doi: https://doi.org/10.2337/dc18-1079.
[14] Gurka MJ, Lilly CL, Oliver MN, DeBoer MD. An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: a confirmatory factor analysis and a resulting continuous severity score. Metabolism 2014;63:218–25. doi: https://doi.org/10.1016/j.metabol.2013.10.006.
[15] Li T, Lei T, Xie Z, Zhang T. Determinants of basic public health services provision by village doctors in China: using non-communicable diseases management as an example. BMC Health Serv Res 2016;16. doi: https://doi.org/10.1186/s12913-016-1276-y.
[16] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMC Med 2015;13:1. doi: https://doi.org/10.1186/s12916-014-0241-z.
[17] Alberti KG et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009;120:1640–5. doi: https://doi.org/10.1161/CIRCULATIONAHA.109.192644.
[18] Liu J et al. Impact of dietary patterns on metabolic syndrome in young adults: a cross-sectional study. Nutrients 2024;16. doi: https://doi.org/10.3390/nu16172890.
[19] Guo H, Jiang W, Zhao B, Xiong Y, Lu Z. A predictive model of metabolic syndrome by medical examination: evidence from an 8-year Chinese cohort. Diabetes Metab Syndr Obes 2021;14:4459–67. doi: https://doi.org/10.2147/DMSO.S314550.
[20] Zou G et al. Predictive analysis of metabolic syndrome based on 5-years continuous physical examination data. Sci Rep 2023;13:9132. doi: https://doi.org/10.1038/s41598-023-35604-8.
[21] Bharat C et al. Development and evaluation of a risk algorithm predicting alcohol dependence after early onset of regular alcohol use. Addiction 2023;118:954–66. doi: https://doi.org/10.1111/add.16122.
[22] Charu V et al. Benchmarking clinical risk prediction algorithms with ensemble machine learning for the noninvasive diagnosis of liver fibrosis in NAFLD. Hepatology 2024. doi: https://doi.org/10.1097/HEP.0000000000000908.

[23] Kanwal F et al. Development, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. JAMA Netw Open 2020;3:e2023780. doi: https://doi.org/10.1001/jamanetworkopen.2020.23780.

[24] Park H et al. Machine learning algorithms for predicting direct-acting antiviral treatment failure in chronic hepatitis C: an HCV-TARGET analysis. Hepatology 2022;76:483–91. doi: https://doi.org/10.1002/hep.32347.

[25] Liao LD et al. Development and validation of prediction models for gestational diabetes treatment modality using supervised machine learning: a population-based cohort study. BMC Med 2022;20:307. doi: https://doi.org/10.1186/s12916-022-02499-7.

[26] Liu W, Tang X, Cui T, Zhao H, Song G. Development and visualization of a risk prediction model for metabolic syndrome: a longitudinal cohort study based on health check-up data in China. Front Nutr 2023;10:1286654. doi: https://doi.org/10.3389/fnut.2023.1286654.

[27] Liu Q et al. Metabolic syndrome parameters' variability and stroke incidence in hypertensive patients: evidence from a functional community cohort. Cardiovasc Diabetol 2024;23:203. doi: https://doi.org/10.1186/s12933-024-02282-3.

[28] Xanthakis VSJ, Samdarshi TE, et al. Relations between subclinical disease markers and type 2 diabetes, metabolic syndrome, and incident cardiovascular disease.pdf. Diabetes Care 2015;38:1082–8. doi: https://doi.org/10.2337/dc14-2460/-/DC1.

[29] Zhang HW et al. Using machine learning to develop a stacking ensemble learning model for the CT radiomics classification of brain metastases. Sci Rep 2024;14:28575. doi: https://doi.org/10.1038/s41598-024-80210-x.

[30] Yom-Tov E et al. Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. J Med Internet Res 2017;19:e338.

[31] Spring B et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013;173:105–11. doi: https://doi.org/10.1001/jamainternmed.2013.1221.

[32] Shah RV et al. Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA Study. J Am Coll Cardiol Img 2014;7:1221–35. doi: https://doi.org/10.1016/j.jcmg.2014.07.017.

[33] van Namen M, Prendergast L, Peiris C. Supervised lifestyle intervention for people with metabolic syndrome improves outcomes and reduces individual risk factors of metabolic syndrome: a systematic review and meta-analysis. Metabolism 2019;101:153988. doi: https://doi.org/10.1016/j.metabol.2019.153988.

[34] Yamaoka K, Tango T. Effects of lifestyle modification on metabolic syndrome: a systematic review and meta-analysis. BMC Med 2012;10:138. doi: https://doi.org/10.1186/1741-7015-10-138.

[35] Guo T et al. The association of long-term trajectories of BMI, its variability, and metabolic syndrome: a 30-year prospective cohort study. EClinicalMedicine 2024;69:102486. doi: https://doi.org/10.1016/j.eclinm.2024.102486.

[36] Marassi M, Fadini GP. The cardio-renal-metabolic connection: a review of the evidence. Cardiovasc Diabetol 2023;22:195. doi: https://doi.org/10.1186/s12933-023-01937-x.

[37] Chan DT, Watts GF, Irish AB, Dogra GK. Insulin resistance and vascular dysfunction in chronic kidney disease: mechanisms and therapeutic interventions. Nephrol Dial Transplant 2017;32:1274–81. doi: https://doi.org/10.1093/ndt/gfv326.

[38] Mauvais-Jarvis F et al. Sex and gender: modifiers of health, disease, and medicine. Lancet 2020;396:565–82. doi: https://doi.org/10.1016/S0140-6736(20)31561-0.

[39] Lonardo A et al. Sex differences in nonalcoholic fatty liver disease: state of the art and identification of research gaps. Hepatology 2019;70:1457–69. doi: https://doi.org/10.1002/hep.30626.

[40] Florio AA et al. Oophorectomy and risk of non-alcoholic fatty liver disease and primary liver cancer in the Clinical Practice Research Datalink. Eur J Epidemiol 2019;34:871–8. doi: https://doi.org/10.1007/s10654-019-00526-1.

[41] Xu F et al. The association between physical activity and metabolic syndrome in older adults with obesity. J Frailty Aging 2019;8:27–32. doi: https://doi.org/10.14283/jfa.2018.34.

[42] Kivimaki M et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. Lancet Public Health 2020;5:e140–9. doi: https://doi.org/10.1016/S2468-2667(19)30248-8.