Original research

Development and Internal Validation of an Interpretable Machine Learning Screening Model for MAFLD–Diabetes Mellitus Comorbidity: A Cross-Sectional NHANES Study

Rui Wang [a#], Yang Cheng [a#], Jie Wang [b], Fengqian Liu [a], Jing Yu [c], Yun Wu [a**]

[a] *Department of General Practice, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830000, Xinjiang, China;*
[b] *Department of Pharmacy, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830000, Xinjiang, China;*
[c] *College of Pharmacy, Xinjiang Medical University, Urumqi 830000, Xinjiang, China.*

ARTICLE INFO

Article history:
Received 09 Jul. 2026
Accepted 16 Sep. 2026

Keywords:
MAFLD;
Diabetes Mellitus;
Risk Prediction Model;
Clinical Decision Support

ABSTRACT

Objective: The rising global prevalence and bidirectional relationship between metabolic-associated fatty liver disease (MAFLD) and diabetes mellitus (DM) significantly exacerbate adverse target-organ outcomes. Despite this, non-invasive screening frameworks remain lacking. This study aims to bridge this gap by developing and validating an individualized, machine learning-based risk prediction model for this comorbidity using routine clinical and biochemical phenotypes. **Methods:** Clinical data of 5,383 adult participants were extracted from the National Health and Nutrition Examination Survey (NHANES 2021-2023) database for retrospective analysis. The cohort was partitioned into a training set and a testing set at a 7:3 ratio. Following feature selection combining LASSO regression and the random forest (RF) algorithm, ten machine learning models were constructed and comparatively evaluated using 10-fold cross-validation. Model discrimination and calibration were comprehensively assessed utilizing the area under the receiver operating characteristic curve (AUC), calibration curves, and Brier scores. The SHAP framework was then applied to quantify feature contributions to the predictive outcomes, culminating in the development of a static web-based tool for individualized risk assessment. **Results:** Compared to the non-comorbidity group (n = 4,801), the comorbidity group (n = 582) exhibited a significantly heavier metabolic burden, characterized by older age and a higher BMI. Following dual feature dimensionality reduction, 23 candidate variables were identified. Among the 10 machine learning algorithms evaluated, the artificial neural network (ANN) demonstrated the best overall performance. Through SHAP-based feature refinement, an optimized ANN model relying on only 13 routine clinical and biochemical indicators was ultimately established. In the testing set, this model achieved excellent discrimination (AUC = 0.863) and the lowest Brier score (0.073), coupled with an exceptionally high probabilistic goodness-of-fit (Hosmer-Lemeshow P = 0.370). Its discriminative performance showed no statistically significant difference from that of the full-feature model (P = 0.894). SHAP explainability analysis indicated that serum osmolality and serum sodium (Na) were the core indicators with the highest predictive contribution to this comorbidity risk. Finally, the online assessment tool deployed from this model successfully enabled the non-invasive, real-time calculation of comorbidity probabilities. **Conclusion:** This study developed and validated a robust 13-feature ANN model for predicting MAFLD-DM comorbidity, highlighting the indicative value of systemic homeostasis and electrolyte imbalances. The derived web-based tool offers cost-effective, non-invasive clinical decision support for early risk stratification and "dual metabolic-hepatic management".

[#] Rui Wang and Yang Cheng are co-first authors. * *Corresponding author: Yun Wu.*
ISSN: 3071-4265/ This work is licensed under a Creative Commons Attribution 4.0 International License.
DOI: 10.64904/20260902

1. Introduction

The overall global prevalence of metabolic dysfunction-associated fatty liver disease (MAFLD) is steadily rising, making it the most common chronic liver disease worldwide [1][2]. In Asian populations, its prevalence ranges from 6% to 25% [3]. Rather than merely a manifestation of localized hepatic lipid deposition, MAFLD is a condition intricately linked to systemic metabolic dysregulation. Its clinical progression spans from simple steatosis and steatohepatitis to fibrosis and, ultimately, cirrhosis, significantly elevating the risk of cardiovascular events, hepatocellular carcinoma (HCC), and various extrahepatic malignancies [4][5][6].

Diabetes mellitus (DM) is one of the major chronic non-communicable diseases threatening global health. With population aging and the increasing prevalence of obesity and sedentary lifestyles, approximately 537 million adults worldwide were living with diabetes in 2021, and this number is projected to continue rising [7][8]. DM has emerged as a leading contributor to the global burden of mortality and disability-adjusted life years (DALYs) [9]. The core pathological hallmark of DM is chronic hyperglycemia resulting from inadequate insulin secretion, insulin resistance, or both. This condition can affect multiple organ systems and trigger a cascade of metabolic and vascular complications [10].

Extensive research indicates that MAFLD and DM frequently coexist. MAFLD increases the risk of developing DM by 1.6- to 6.8-fold [11], and the prevalence of MAFLD among patients with type 2 diabetes mellitus (T2DM) reaches up to 75% [12]. These two conditions exhibit profound mechanistic overlap. A complex interplay of insulin resistance, lipotoxicity, chronic low-grade inflammation, oxidative stress, intestinal dysbiosis, and genetic predisposition synergistically drives their pathogenesis and progression [13][14]. Clinically, DM not only significantly accelerates the progression of MAFLD to steatohepatitis, advanced fibrosis, and ultimately end-stage liver disease, but also elevates the risk of HCC [15][16]. Conversely, the presence of MAFLD exacerbates the metabolic burden in individuals with DM, increasing the incidence and progression of diabetes-related complications such as cardiovascular disease and chronic kidney disease, thereby establishing a vicious cycle [17][18]. Consequently, this interplay poses a formidable challenge to global public health.

Despite the well-established high comorbidity rate and bidirectional synergistic effects of MAFLD and DM, physicians and patients in real-world clinical practice often focus on a single disease. This siloed approach frequently leads to under-screening, fragmented treatment, and delayed management, thereby exacerbating the burden of healthcare and chronic disease management [19]. Against this backdrop, the paradigm of 'integrated diabetes-liver co-management' has emerged. Its objective is to achieve the unified management of comorbidities and their shared etiologies through multidisciplinary collaboration, yielding simultaneous benefits for glycemic and hepatic outcomes within a single intervention framework [20]. However, a critical prerequisite for this co-management model is identifying the susceptibility characteristics and risk factors for DM with comorbid MAFLD at the population level, and establishing a simple, reliable risk identification tool to guide screening and intervention for susceptible populations. Current studies on this comorbidity remain limited. Previous evidence is mostly derived from single-center cohorts or specific populations with insufficient variable coverage, failing to reflect the true exposure structure and risk disparities in the general population. Furthermore, methodologically, traditional linear models often struggle to achieve precise fitting and prediction when dealing with high-dimensional clinical data, multicollinearity, and complex non-linear associations inherent in complex metabolic networks. In recent years, machine learning (ML) has demonstrated tremendous advantages in complex comorbidity risk assessment due to its exceptional capacity for high-dimensional data mining and non-linear feature extraction [21][22]. Nevertheless, large-population studies comprehensively comparing various ML algorithms to construct risk prediction models for DM and MAFLD remain scarce, and there is a conspicuous lack of intelligent assessment tools ready for direct clinical deployment.

In light of this, the present study conducted a retrospective analysis using the National Health and Nutrition Examination Survey (NHANES) database, a highly nationally representative dataset characterized by multi-ethnic coverage and comprehensive clinical parameters. We aimed to comprehensively explore the exposure characteristics and associated risk factors in patients with comorbid DM and MAFLD. By employing various machine learning algorithms, we sought to select the optimal core features and construct a robust clinical risk prediction model for this comorbidity. Furthermore, to overcome barriers to clinical translation, we transformed the optimal model into an interactive web-based assessment tool. Ultimately, this study aims to provide solid theoretical evidence and a pragmatic instrument for the early precision screening, risk stratification of susceptible populations, and the comprehensive implementation of the “integrated diabetes-liver co-management” strategy.

2. Methods

2.1. Study Design and Data Source

This cross-sectional study utilized data from the NHANES spanning the 2021 – 2023 cycle. Conducted by the National Center for Health Statistics (NCHS), NHANES is a nationally representative survey that employs a complex, multistage probability design to assess the health and nutritional status of the U.S. population. The comprehensive data collection encompasses demographic characteristics, physical examinations, laboratory measurements, and disease status questionnaires [23]. The NHANES study protocols were approved by the NCHS Research Ethics Review Board, and written informed consent was obtained from all participants.

2.2. Study Population

Initially, a total of 11,933 participants from the NHANES 2021 – 2023 cycle were identified. To ensure the rigor of the analysis, participants were sequentially excluded based on the following pre-established criteria: (1) missing liver ultrasound examination data (n = 5,520); (2) lacking clinical diagnostic information for diabetes mellitus (n = 193); and (3) aged < 20 years (n = 837). Following these exclusions, a final cohort of 5,383 eligible participants was included in the current analysis. The detailed participant selection flowchart is illustrated in Figure 1.

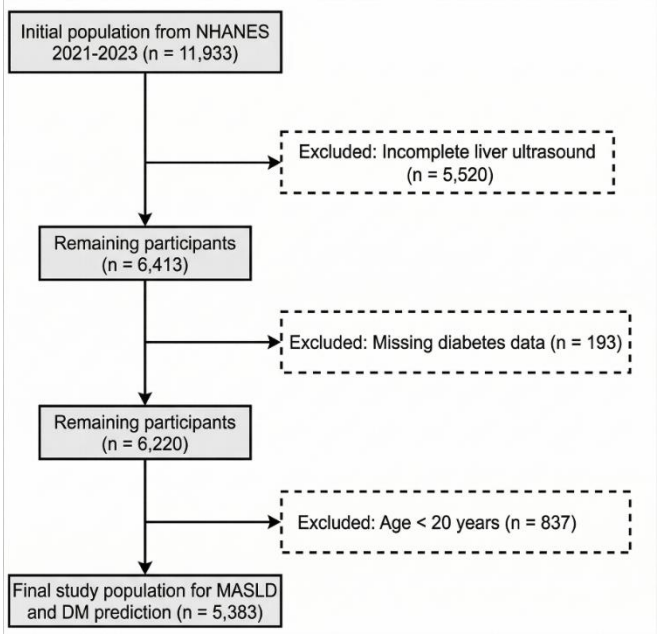

Figure 1: Flowchart of participant selection

2.3. Definitions of MAFLD

Following international consensus guidelines, MAFLD was diagnosed based on hepatic steatosis with metabolic abnormalities. Hepatic steatosis was diagnosed exclusively using vibration-controlled transient elastography with Controlled Attenuation Parameter ⩾ 248 dB/m, which represents the only objective hepatic steatosis assessment available in NHANES 2017-2023. This threshold has been validated in large meta-analyses for detecting hepatic steatosis with a sensitivity of 69% and specificity of 82% for any grade of steatosis [24].

MAFLD diagnosis required at least one of the following criteria: overweight/obesity (body mass index (BMI) ⩾25 kg/m^2), type 2 diabetes, or evidence of metabolic dysregulation. Metabolic dysregulation was defined by the presence of at least two of the following: waist circumference ⩾102 cm in men or ⩾88 cm in women, blood pressure ⩾ 130/85 mmHg or antihypertensive medication use, triglycerides ⩾1.70 mmol/L or lipid-lowering therapy, HDL cholesterol <1.0 mmol/L in men or <1.3 mmol/L in women, prediabetes, Homeostatic Model Assessment of Insulin Resistance (HOMA-IR) ⩾ 2.5, or high-sensitivity C-reactive protein >2 mg/L [25][26].

2.4. Definitions of DM

Diabetes was defined as HbA1c > 6.5%, fasting glucose ≥ 7.0 mmol/L, random blood glucose or 2-h Oral Glucose Tolerance Test (OGTT) blood glucose ≥ 11.1 mmol/L, use of diabetes medications or a previous diagnosis of diabetes [27].

2.5. Study Outcomes and Cohort Categorization

The primary outcome of this study was defined as the comorbidity of MAFLD and DM. Based on whether they simultaneously met the clinical diagnostic criteria for both conditions, the study participants were categorized into two groups for comparative analysis: the comorbidity group (concurrent MAFLD and DM) and the non-comorbidity group (absence of concurrent MAFLD and DM).

2.6. Candidate Predictors and Feature Engineering

To construct a comprehensive and robust predictive model, a wide array of baseline variables was extracted from the NHANES database to serve as candidate predictors. These variables encompassed sociodemographic characteristics, lifestyle factors, anthropometric measurements, comprehensive laboratory parameters, as well as various inflammatory markers and environmental exposures.

2.7. Statistical Analysis and Feature Selection

To identify independent clinical factors associated with concurrent MAFLD and DM, candidate variables were first screened using univariate logistic regression, and those with $P < 0.05$ were included in a multivariable model to calculate adjusted odds ratios (ORs) and 95% confidence intervals (CIs). Furthermore, given the high dimensionality of the candidate predictors, a hybrid feature selection strategy was implemented for subsequent machine learning. First, LASSO regression with 10-fold cross-validation was utilized for dimensionality reduction. Concurrently, a Random Forest (RF) algorithm was applied to evaluate and rank feature importance based on the Mean Decrease Gini (MDG) index. To maximize model robustness, only features identified by both LASSO and RF (the intersection set) were retained as the optimal subset for machine learning model development.

2.8. Machine Learning Model Development and Validation

Ten machine learning algorithms were evaluated to predict concurrent MAFLD and DM: Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Extra Trees (ET), Gradient Boosting Machine (GBM), AdaBoost, XGBoost, and Artificial Neural Network (ANN). The total study cohort

was partitioned into a training set and a testing set using a stratified random sampling approach (ratio 7:3) to ensure that the prevalence of the outcome remained consistent across both subsets. Within the training set, a 10-fold cross-validation strategy was employed for hyperparameter optimization and to mitigate potential overfitting.

Recursive Feature Elimination (RFE) was applied to the best-performing algorithm. By assessing the impact of varying feature counts on the Area Under the Receiver Operating Characteristic Curve (AUC), the optimal feature subset was identified at the inflection point where additional variables yielded no significant predictive gain. Subsequently, model performance in the independent testing set was evaluated using discrimination metrics (AUC, accuracy, sensitivity, specificity, precision, recall, and F1-score) and calibration measures (calibration curves and Brier score).

2.9. Model Interpretability With SHAP

The Shapley Additive Explanations (SHAP) framework was employed to interpret the predictions of the best-performing model. By calculating the average marginal contribution of each feature across all possible coalitions, SHAP provides both global feature importance and local instance-level explanations. SHAP summary plots and feature importance rankings were generated to visualize the magnitude and directionality of each variable's impact on predicting concurrent MAFLD and DM.

2.10. Development of the Web-Based Calculator

To facilitate clinical translation, the best-performing predictive model was deployed as an interactive, web-based calculator. This open-access tool allows clinicians to input the values of the finalized optimal features and dynamically generates an individualized probability of concurrent MAFLD and DM in real-time. The deployment of this calculator bridges the gap between complex machine learning algorithms and point-of-care clinical decision-making, thereby significantly enhancing the model's accessibility and practical utility.

2.11. Statistical Analyses

The normality of continuous variables was evaluated using the Shapiro-Wilk test. Normally distributed continuous variables were expressed as mean±standard deviation (SD) and compared using the independent samples t-test. Non-normally distributed variables were presented as medians with interquartile ranges (IQRs) and compared using the Mann-Whitney U test. Categorical variables were presented as frequencies and percentages, and group differences were assessed utilizing the Chi-square test or Fisher's exact test, as appropriate. Before machine learning modeling, all continuous features were standardized using Z-score normalization to ensure uniform scaling. All computational and statistical analyses were performed using R software (version 4.4.3). A two-sided P-value < 0.05 was considered to indicate statistical significance.

3. Results

3.1. Study Population and Selection

A total of 11,933 participants were initially identified from the database. Following the sequential exclusion of individuals with incomplete liver ultrasound data (n = 5,520), missing diabetes-related diagnostic information (n = 193), and those aged < 20 years (n = 837), a final cohort of 5,383 eligible participants was included in the study. The detailed participant selection process is illustrated in Figure 1.

3.2. Baseline Characteristics

The final study cohort (n = 5,383) was stratified into a comorbidity group (concurrent MAFLD and DM, n = 582) and a non-comorbidity group (n = 4,801). Table 1 summarizes the core demographic, clinical, and primary laboratory characteristics of the participants. Compared to the non-comorbidity group, individuals in the comorbidity group were significantly older and had a higher body mass index (BMI) (both P < 0.001). Furthermore, the comorbidity group exhibited profound alterations in glucose and lipid metabolism, liver enzymes, multiple electrolytes (e.g., sodium and magnesium), and serum osmolality. A comprehensive and exhaustive comparison of the extended baseline profiles, including complete blood counts, inflammation derived indices, trace elements, and endocrine hormones, is detailed in Supplementary Tables S1–S6.

Table 1. Core baseline demographic and clinical characteristics of the study population

Variables	Total (N = 5,383)	Non-comorbidity grou (n = 4,801)	Comorbidity group (n = 582)	*P*
Age, years	57.00 (39.00, 68.00)	55.00 (38.00, 67.00)	63.00 (55.00, 70.00)	<0.001
Gender, n (%)				0.136
Female	2931 (54.45)	2631 (54.80)	300 (51.55)	
Male	2452 (45.55)	2170 (45.20)	282 (48.45)	
Race, n (%)				0.229
Mexican American	368 (6.84)	315 (6.56)	53 (9.11)	
Non-Hispanic Black	643 (11.95)	571 (11.89)	72 (12.37)	

Variables	Total (N = 5,383)	Non-comorbidity grou (n = 4,801)	Comorbidity group (n = 582)	*P*
Non-Hispanic White	3170 (58.89)	2839 (59.13)	331 (56.87)	
Other Hispanic	541 (10.05)	485 (10.10)	56 (9.62)	
Other Race	661 (12.28)	591 (12.31)	70 (12.03)	
Marital status, n (%)				<0.001
Married/Living with partner	2948 (54.77)	2633 (54.84)	315 (54.12)	
Never married	1110 (20.62)	1031 (21.47)	79 (13.57)	
Widowed/Divorced/Separated	1325 (24.61)	1137 (23.68)	188 (32.30)	
Education level, n (%)				<0.001
< High school	662 (12.30)	545 (11.35)	117 (20.10)	
High school	1132 (21.03)	991 (20.64)	141 (24.23)	
> High school	3589 (66.67)	3265 (68.01)	324 (55.67)	
Smoking status, n (%)				<0.001
No	3161 (58.72)	2869 (59.76)	292 (50.17)	
Yes	2222 (41.28)	1932 (40.24)	290 (49.83)	
Drinking status, n (%)				0.005
No	632 (11.74)	543 (11.31)	89 (15.29)	
Yes	4751 (88.26)	4258 (88.69)	493 (84.71)	
BMI, kg/m^2	28.30 (24.60, 33.30)	27.70 (24.30, 32.20)	34.45 (30.30, 39.48)	<0.001
MAFLD, n (%)				<0.001
No	3501 (65.04)	3501 (72.92)	0 (0.00)	
Yes	1882 (34.96)	1300 (27.08)	582 (100.00)	
TC, mmol/L	4.81 (4.14, 5.56)	4.84 (4.16, 5.56)	4.52 (3.80, 5.30)	<0.001
Triglycerides, mmol/L	1.24 (0.89, 1.80)	1.20 (0.86, 1.70)	1.80 (1.27, 2.62)	<0.001
HDL-C, mmol/L	1.34 (1.14, 1.60)	1.40 (1.16, 1.66)	1.14 (1.01, 1.34)	<0.001
FBG, mmol/L	5.00 (5.00, 6.00)	5.00 (5.00, 6.00)	7.00 (6.00, 9.00)	<0.001
HbA1c, %	5.50 (5.20, 5.90)	5.50 (5.20, 5.80)	6.90 (6.30, 8.00)	<0.001
SCR, M (Q_1, Q_3)	75.14 (63.65, 88.40)	75.14 (63.65, 87.52)	75.14 (62.76, 91.72)	0.903
SALT, U/L	18.00 (13.00, 26.00)	17.00 (13.00, 25.00)	23.00 (16.00, 32.00)	<0.001
GGT, U/L	20.00 (14.00, 32.00)	19.00 (13.00, 30.00)	28.50 (19.00, 46.00)	<0.001
Osmolality, mOsm/kg	280.00 (276.00, 283.00)	279.00 (276.00, 282.00)	282.00 (279.00, 285.00)	<0.001
Mg, mg/dL	2.00 (1.90, 2.10)	2.00 (1.90, 2.10)	1.90 (1.70, 2.00)	<0.001
Na, mmol/L	140.00 (138.00, 141.00)	140.00 (138.00, 141.00)	139.00 (138.00, 141.00)	<0.001
Cl, mmol/L	102.00 (100.00, 103.00)	102.00 (100.00, 103.00)	101.00 (99.00, 103.00)	<0.001
RBC, $10^{12}/L$	4.65 (4.36, 4.95)	4.64 (4.35, 4.94)	4.76 (4.45, 5.07)	<0.001
WBC, $10^9/L$	6.60 (5.50, 7.90)	6.50 (5.40, 7.80)	7.60 (6.20, 8.90)	<0.001
Neutrophil, $10^9/L$	3.90 (3.00, 4.90)	3.80 (3.00, 4.80)	4.40 (3.50, 5.50)	<0.001
LYMNO, $10^9/L$	1.90 (1.50, 2.40)	1.90 (1.50, 2.40)	2.10 (1.70, 2.70)	<0.001
MCV, fL	89.10 (85.90, 92.10)	89.30 (86.10, 92.30)	87.90 (84.60, 91.00)	<0.001
RDW, %	13.60 (13.10, 14.20)	13.50 (13.00, 14.10)	14.00 (13.40, 14.70)	<0.001

Variables	Total (N = 5,383)	Non-comorbidity grou (n = 4,801)	Comorbidity group (n = 582)	*P*
Testosterone, nmol/L	1.57 (0.68, 14.30)	1.56 (0.68, 14.70)	1.58 (0.58, 10.97)	<0.001
SHBG, nmol/L	43.83 (29.73, 64.90)	45.26 (30.90, 66.60)	33.39 (24.30, 47.19)	<0.001
DHEAS, μmol/L	1.99 (1.02, 3.63)	2.07 (1.07, 3.77)	1.44 (0.73, 2.50)	<0.001
Androstenedione, nmol/L	2.15 (1.53, 3.09)	2.19 (1.55, 3.17)	1.85 (1.31, 2.58)	<0.001
FSH, IU/L	8.20 (4.50, 46.86)	7.82 (4.32, 47.31)	15.25 (6.46, 43.85)	<0.001
VD_3, nmol/L	73.20 (49.95, 97.30)	73.60 (51.00, 97.30)	69.45 (42.60, 95.92)	0.006
Pb, μmol/L	0.04 (0.02, 0.06)	0.04 (0.02, 0.06)	0.03 (0.02, 0.05)	<0.001
Se, μmol/L	2.26 (2.09, 2.45)	2.26 (2.09, 2.45)	2.29 (2.12, 2.49)	0.078
Mn, nmol/L	160.72 (129.51, 200.59)	159.81 (129.23, 199.86)	167.19 (132.87, 206.37)	0.028

Note: Data are presented as median (interquartile range, Q1, Q3) for continuous variables due to their non-normal distribution, or as number (percentage) for categorical variables. *P* values were calculated using the Mann-Whitney U test for continuous variables and the Chi-square test for categorical variables. BMI, body mass index; MAFLD, metabolic dysfunction-associated fatty liver disease; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; FBG, fasting blood glucose; HbA1c, glycated hemoglobin; SCR, serum creatinine; SALT, serum alanine aminotransferase; GGT, gamma-glutamyl transferase; Mg, magnesium; Na, sodium; Cl, chloride; RBC, red blood cell; WBC, white blood cell; MCV, mean corpuscular volume; RDW, red cell distribution width; SHBG, sex hormone-binding globulin; DHEAS, dehydroepiandrosterone sulfate; FSH, follicle-stimulating hormone; VD_3, 25-hydroxyvitamin D_3; Pb, lead; Se, selenium; Mn, manganese; Osmolality, serum osmolality; LYMNO, Lymphocyte.

3.3. Feature Selection

The LASSO regression model was applied to penalize and select candidate predictors. Based on the optimal tuning parameter (λ), 41 variables were retained (Figure 2). Concurrently, the RF algorithm evaluated feature importance, identifying top-ranking variables based on a Mean Decrease Gini (MDG) threshold of ≥ 13.13 (Figure 3). By taking the intersection of the features identified by both algorithms, a final subset of 23 optimal variables was established for subsequent model development (Figure 4). These 23 core features included: age, red blood cell (RBC) count, white blood cell (WBC) count, lymphocyte count, mean corpuscular volume (MCV), red blood cell distribution width (RDW), sodium (Na), chloride (Cl), serum osmolality, total cholesterol (TC), serum creatinine (SCR), magnesium (Mg), manganese (Mn), lead (Pb), selenium (Se), androstenedione, dehydroepiandrosterone sulfate (DHEAS), follicle-stimulating hormone (FSH), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), sex hormone-binding globulin (SHBG), testosterone, and VD3.

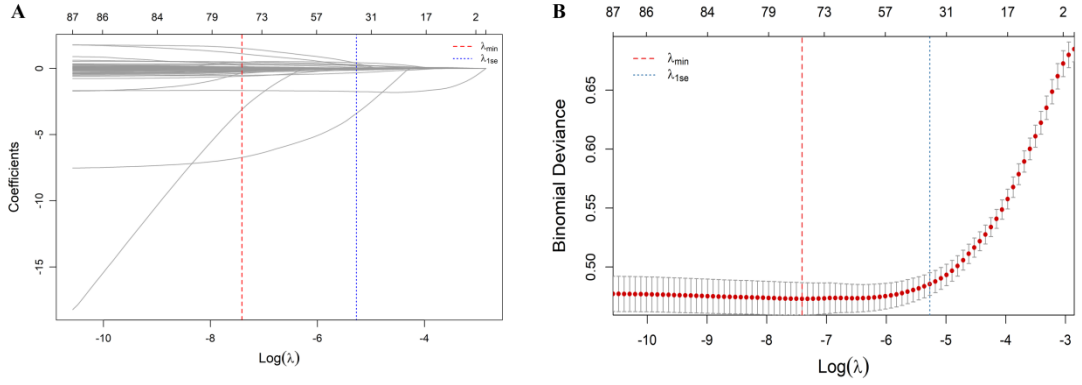

Figure 2: Feature selection using the LASSO logistic regression model

Note: (A) LASSO coefficient profiles of the candidate variables plotted against log(λ). Each curve represents the trajectory of a single variable's coefficient. The top axis indicates the number of variables with non-zero coefficients at the corresponding λ value. (B) Selection of the optimal tuning parameter (λ) via 10-fold cross-validation. The red dots and error bars represent the mean binomial deviance and corresponding standard errors at different log(λ) values. The vertical red dashed line indicates the λ value that minimizes the cross-validation error (λ_{min}), while the vertical blue dotted line represents the largest λ value within one standard error of the minimum (λ_{1se}).

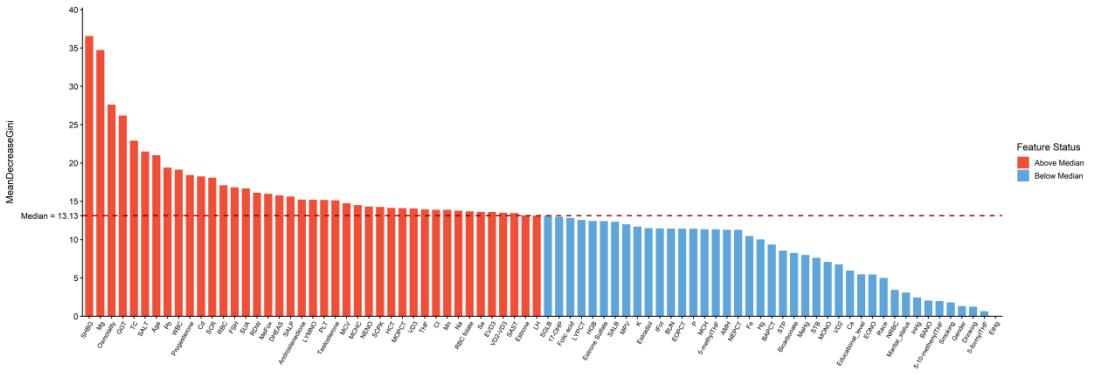

Figure 3: Feature importance ranking and selection based on the RF algorithm

Note: Variables are displayed on the x-axis in descending order of their MDG values. The horizontal dashed line indicates the selection threshold, which was set at the median MDG value of 13.13. Features with an MDG ≥ 13.13 (red bars) were identified as having a highly significant contribution to the model's predictive perf

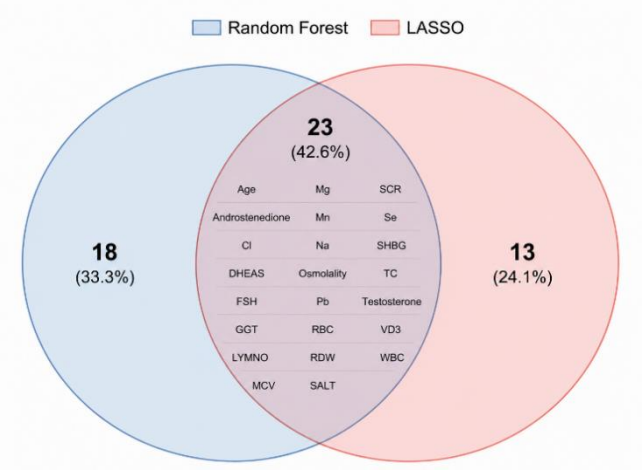

Figure 4: Venn diagram of the optimal feature subset selected by LASSO regression and the RF algorithm

Note: The blue circle represents the features selected by the RF algorithm, while the red circle represents the features penalized and retained by the LASSO regression. The overlapping central region indicates the intersection of the two feature sets, comprising 23 core predictive variables (42.6% of the total combined features) jointly identified by both methods.

3.4. Univariate and Multivariable Logistic Regression Analysis for the Comorbidity of MAFLD and DM

A univariable logistic regression analysis was conducted on the 23 pre-selected candidate variables to evaluate their unadjusted associations with the risk of comorbid MAFLD and DM. 20 variables exhibited significant statistical associations with the outcome ($P < 0.05$). Specifically, higher levels of age, ALT, GGT, RBC, WBC, lymphocytes, RDW, serum osmolality, and Mn were associated with an increased risk. Conversely, TC, Mg, Na, Cl, MCV, testosterone, SHBG, DHEAS, androstenedione, VD_3, and Pb were inversely associated with the risk. No significant associations were observed for SCR, FSH, and Se ($P > 0.05$).

Variables identified as significant in the univariate analysis were subsequently entered into a multivariable logistic regression model to identify independent predictors. The results revealed that age, GGT, ALT, serum osmolality, RBC, WBC, lymphocyte, and Mn remained independent risk factors for comorbid MAFLD and DM ($P < 0.05$). However, TC, Na, Cl, Mg, testosterone, SHBG, DHEAS, VD3, and Pb were identified as independent protective factors ($P < 0.05$). Notably, serum osmolality (OR = 1.260, 95% CI: 1.220-1.300) and RBC (OR = 1.971, 95% CI: 1.509-2.575) showed strong positive associations, while Mg (OR = 0.103, 95% CI: 0.061-0.174) showed a potent protective effect **(Table 2)**.

Table 2. Univariate and multivariable logistic regression analysis of factors associated with the comorbidity of MAFLD and DM

Variables	Univariate Logistic	Multivariable Logistic

	OR (95%CI)	*P*	OR (95%CI)	*P*
Age	1.032(1.026-1.038)	＜0.001	1.038(1.028-1.049)	＜0.001
TC	0.794(0.731-0.863)	＜0.001	0.808(0.734-0.889)	＜0.001
SCR	1.001(0.999-1.003)	0.552	—	
ALT	1.015(1.011-1.019)	＜0.001	1.012(1.007-1.017)	＜0.001
GGT	1.005(1.004-1.007)	＜0.001	1.005(1.003-1.006)	＜0.001
Serum osmolality	1.105(1.086-1.124)	＜0.001	1.260(1.220-1.300)	＜0.001
Mg	0.050(0.032-0.078)	＜0.001	0.103(0.061-0.174)	＜0.001
Na	0.938(0.906-0.971)	＜0.001	0.676(0.629-0.727)	＜0.001
Cl	0.900(0.876-0.925)	＜0.001	0.960(0.922-0.100)	0.050
RBC	1.641(1.373-1.962)	＜0.001	1.971(1.509-2.575)	＜0.001
WBC	1.230(1.184-1.278)	＜0.001	1.067(1.008-1.131)	0.026
Lymphocyte	1.545(1.386-1.723)	＜0.001	1.390(1.192-1.621)	＜0.001
MCV	0.957(0.944-0.970)	＜0.001	0.981(0.959-1.003)	0.089
RDW	1.184(1.127-1.244)	＜0.001	1.075(0.999-1.157)	0.054
Testosterone	0.978(0.968-0.988)	＜0.001	0.970(0.954-0.986)	＜0.001
SHBG	0.978(0.974-0.983)	＜0.001	0.979(0.974-0.984)	＜0.001
DHEAS	0.786(0.744-0.830)	＜0.001	0.880(0.812-0.954)	0.002
Androstenedione	0.728(0.673-0.787)	＜0.001	0.950(0.861-1.048)	0.306
FSH	1.001(0.100-1.003)	0.584	—	
VD3	0.997(0.995-0.100)	0.034	0.994(0.991-0.997)	＜0.001
Pb	0.012(0.001-0.268)	0.003	0.000(0.000-0.007)	＜0.001
Se	1.083(0.881-1.332)	0.448	—	
Mn	1.002(1.000-1.003)	0.008	1.002(1.001-1.004)	0.007

3.5. Correlation Analysis of Clinical Characteristics

Spearman correlation analysis was conducted to evaluate the monotonic relationships among the 23 candidate clinical variables and their associations with the risk of comorbid MAFLD and DM (Figure 5). The correlation matrix revealed that GGT (r=0.19), serum osmolality (r=0.18), WBC (r=0.15), and RDW (r=0.15) exhibited the strongest positive correlations with the comorbidity outcome. Conversely, SHBG (r=-0.17), Mg (r=-0.16), and DHEAS (r=-0.12) demonstrated notable negative correlations. Furthermore, the heatmap identified substantial multicollinearity among specific baseline predictors. For instance, a strong positive correlation was observed between serum osmolality and Na (r=0.81), while age and DHEAS showed a strong negative correlation (r=-0.65).

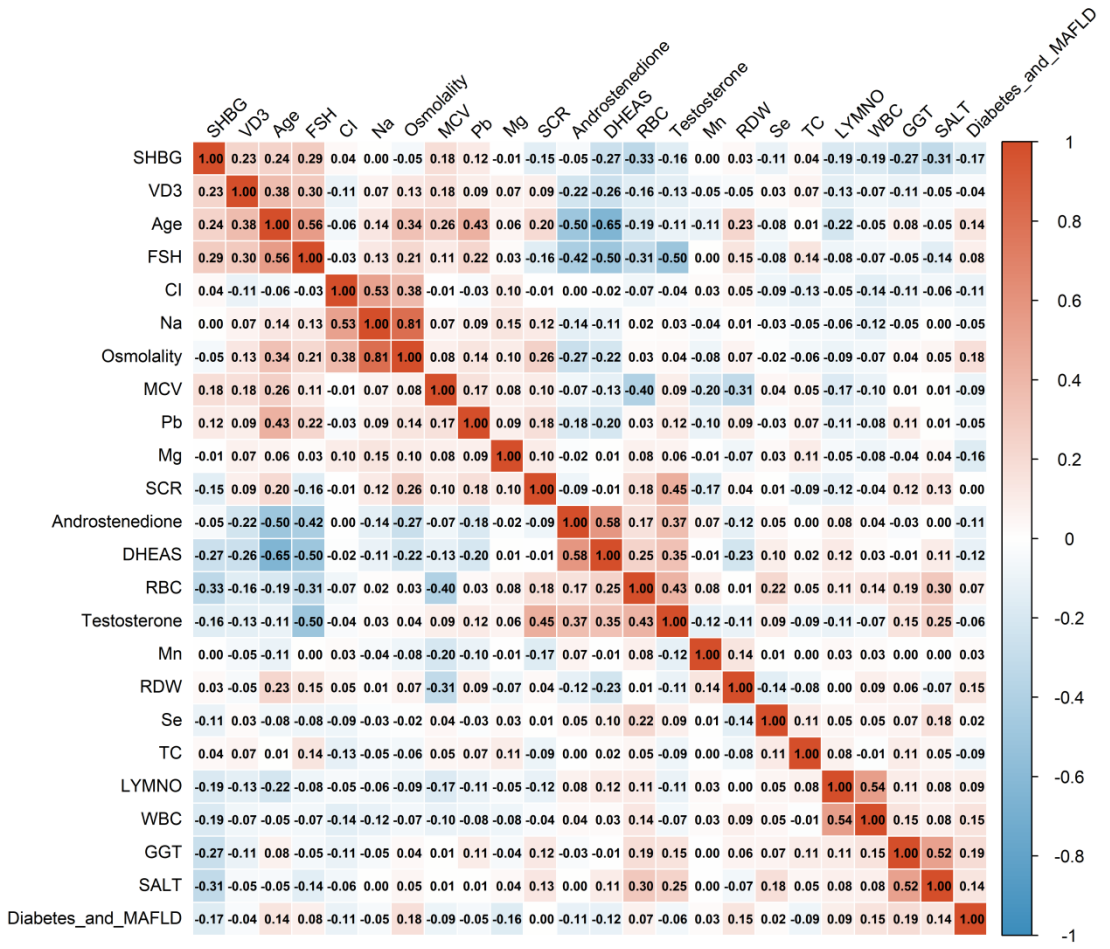

Figure 5: The correlation heat map is used to determine the relationship between variables.

Note: Colors represent the direction and magnitude of the Spearman correlation coefficients (r). Red shades indicate positive correlations, while blue shades indicate negative correlations. The intensity of the color and the numeric value within each cell correspond to the strength of the correlation. DM, diabetes mellitus; MAFLD, metabolic dysfunction-associated fatty liver disease; SHBG, sex hormone-binding globulin; VD3, vitamin D3; FSH, follicle-stimulating hormone; Cl, chloride; Na, sodium; MCV, mean corpuscular volume; Pb, lead; Mg, magnesium; SCR, serum creatinine; DHEAS, dehydroepiandrosterone sulfate; RBC, red blood cell; Mn, manganese; RDW, red blood cell distribution width; Se, selenium; TC, total cholesterol; LYMNO, lymphocyte count; WBC, white blood cell; GGT, gamma-glutamyl transferase; SALT, serum alanine aminotransferase.

3.6. Development and Validation of Machine Learning Models for Predicting the Comorbidity Risk of MAFLD and DM

Based on the selected feature variables, risk prediction models were constructed using ten machine learning algorithms: LR, SVM, DT, RF, KNN, ET, GBM, AdaBoost, XGBoost, and ANN. The comprehensive performance of each model across metrics, including accuracy, sensitivity, specificity, precision, recall, and F1 score, is visualized in the heatmap (Figure 6A).

To evaluate model discrimination, receiver operating characteristic (ROC) curves were plotted, and the AUC was calculated for both the training and test sets. The results indicated that tree-based models, such as ET and RF, achieved exceptionally high AUC values in the training set (0.9537 and 0.9085, respectively). However, their AUC values showed significant attenuation on the test set (to 0.8528 and 0.8215, respectively), suggesting some overfitting. In contrast, the ANN model demonstrated superior robustness and generalization capability; it not only maintained high predictive performance in the training set (AUC = 0.8911) but also achieved the highest AUC among all models in the test set (0.8632), indicating its optimal performance in identifying the comorbidity risk of MAFLD and diabetes (Figure 6B and 6C).

Regarding model calibration and goodness-of-fit in the test set, the ANN model exhibited a distinct advantage. Visual assessment of the calibration curves revealed that the predicted probability curve of the ANN model was most closely aligned with the ideal reference line (Figure 6D). Quantitative

analysis further confirmed that the ANN model yielded the lowest Brier score of 0.0735 (95% CI: 0.0647-0.0832). Crucially, the Hosmer-Lemeshow (H-L) test results demonstrated that, among the ten evaluated models, the ANN model was the only one that showed no statistically significant difference between predicted probabilities and actual incidence rates (χ^2 = 3.1414, P = 0.370), suggesting excellent goodness-of-fit. Conversely, the remaining nine algorithms all yielded statistically significant H-L test results ($P < 0.05$), with the AdaBoost model exhibiting the poorest calibration performance, characterized by the highest Brier score (0.1142, 95% CI: 0.1089-0.1195) and the most significant deviation in the H-L test ($P < 0.001$) (Table 3). Taken together, the assessments of discrimination, calibration curves, and quantitative metrics identify the ANN model as the ideal predictive model, offering the highest accuracy and probability consistency.

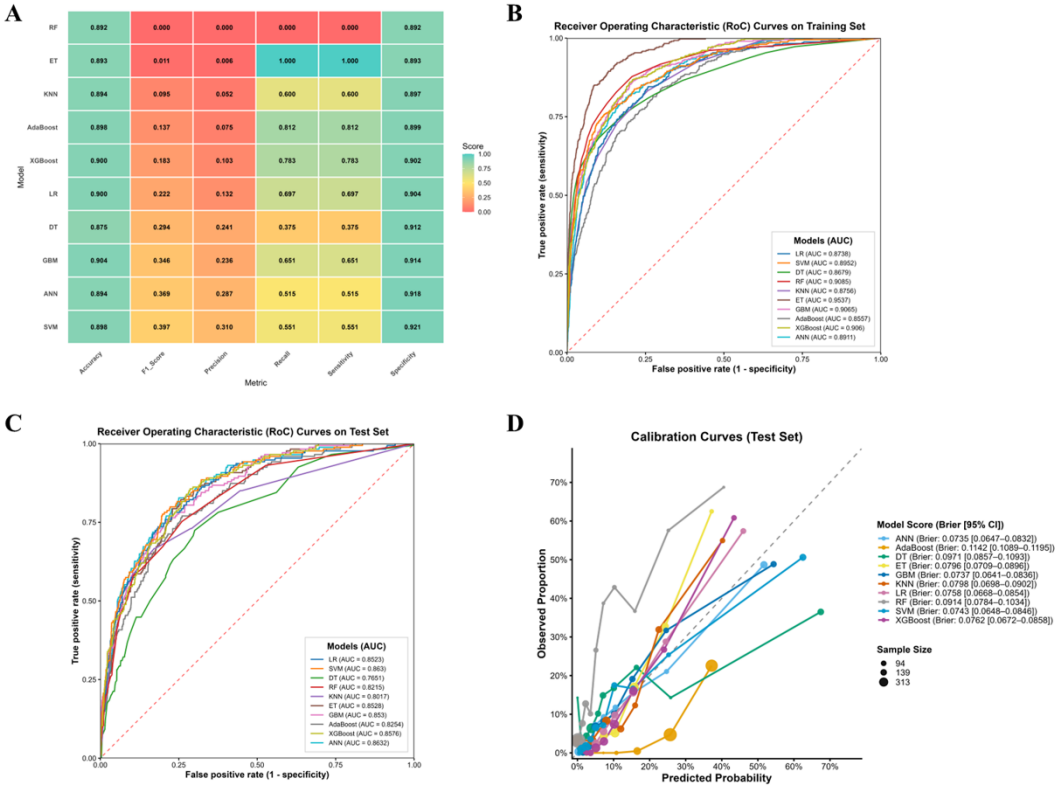

Figure 6: Comprehensive performance, discrimination, and calibration evaluation of the ten machine learning models

Note: (A) Heatmap of multiple performance metrics (Accuracy, F1 Score, Precision, Recall, Sensitivity, and Specificity) across different models. (B) Receiver Operating Characteristic (ROC) curves indicating the discrimination performance of the models on the training set. (C) ROC curves on the test set. (D) Calibration curves and Brier scores of the models on the test set. The diagonal dashed line represents perfect calibration. Point sizes indicate the sample size within each predicted probability bin. RF, Random Forest; ET, Extra Trees; KNN, K-Nearest Neighbors; AdaBoost, Adaptive Boosting; XGBoost, Extreme Gradient Boosting; LR, Logistic Regression; DT, Decision Tree; GBM, Gradient Boosting Machine; ANN, Artificial Neural Network; SVM, Support Vector Machine; AUC, Area Under the Curve; CI, Confidence Interval.

Table 3. Hosmer-Lemeshow goodness-of-fit test for the ten machine learning models

Model	χ^2	P
DT	94.2892	<0.001
RF	789.4795	<0.001
KNN	33.9973	<0.001
ET	51.1842	<0.001
AdaBoost	286.7362	<0.001
XGBoost	39.9003	<0.001
LR	20.3425	0.005
SVM	13.569	0.009
GBM	8.825	0.032
ANN	3.1414	0.370

Note: χ^2, Chi-square statistic from the Hosmer-Lemeshow test. A P-value > 0.05 indicates an acceptable calibration and a good fit between the predicted probabilities and observed outcomes. ANN, artificial neural network; DT, decision tree; ET, extra trees; GBM, gradient boosting machine; KNN, k-nearest neighbors; LR, logistic regression; RF, random forest; SVM, support vector machine.

3.7. SHAP-based Selection of Optimal Predictors and Model Interpretability Analysis

To enhance clinical utility while maintaining predictive performance, a backward elimination strategy based on SHAP feature importance rankings was employed to screen the 17 initial variables. During the progressive reduction of features, the AUC of the ANN model remained broadly stable, peaking at 13 core features (AUC = 0.864). Further reduction below 13 features led to a distinct decline in the AUC, indicating that this 13-feature subset represents the optimal balance between model performance and parsimony (Figure 7A).

The robustness of this feature subset was further validated using Delong's nonparametric test (Figure 7B). Compared with the full 17-feature baseline model (AUC = 0.8632), reducing the model to 14 features (ΔAUC = 0.0090, P = 0.093) or 13 features (ΔAUC = 0.0010, P = 0.894) resulted in no significant loss of discrimination. However, reducing the feature count to 12 or fewer significantly impaired the discriminative capability (e.g., for 12 features: ΔAUC = 0.0190, P = 0.002). Consequently, the 13-feature combination was confirmed as the optimal predictive subset for the ANN model.

The SHAP method was subsequently applied to interpret the logic of this optimized model (Figure 7C). Quantitative evaluation of feature importance revealed that serum osmolality and Na were the strongest contributors to the ANN predictions, with mean absolute SHAP values of 0.089 and 0.078, respectively. These were followed by SHBG (0.034), Mg (0.030), GGT (0.030), and age (0.030). The remaining variables also demonstrated varying degrees of predictive value.

Furthermore, the SHAP summary plot illustrated the directional impact of varying feature values on comorbidity risk (Figure 7D). Serum osmolality and Na exhibited the widest dispersion of SHAP values, indicating that inter-individual variances in these parameters heavily drive model discrimination. Additionally, higher GGT levels were predominantly clustered in the positive SHAP value region, suggesting a positive association with increased comorbidity risk. Conversely, higher SHBG levels were concentrated in the negative SHAP region, implying a potential protective effect against the outcome. Collectively, these results demonstrate that the ANN model's predictive logic relies predominantly on patient osmolality, electrolyte balance, and metabolic status, thereby providing a robust, biologically plausible rationale for the model's outputs.

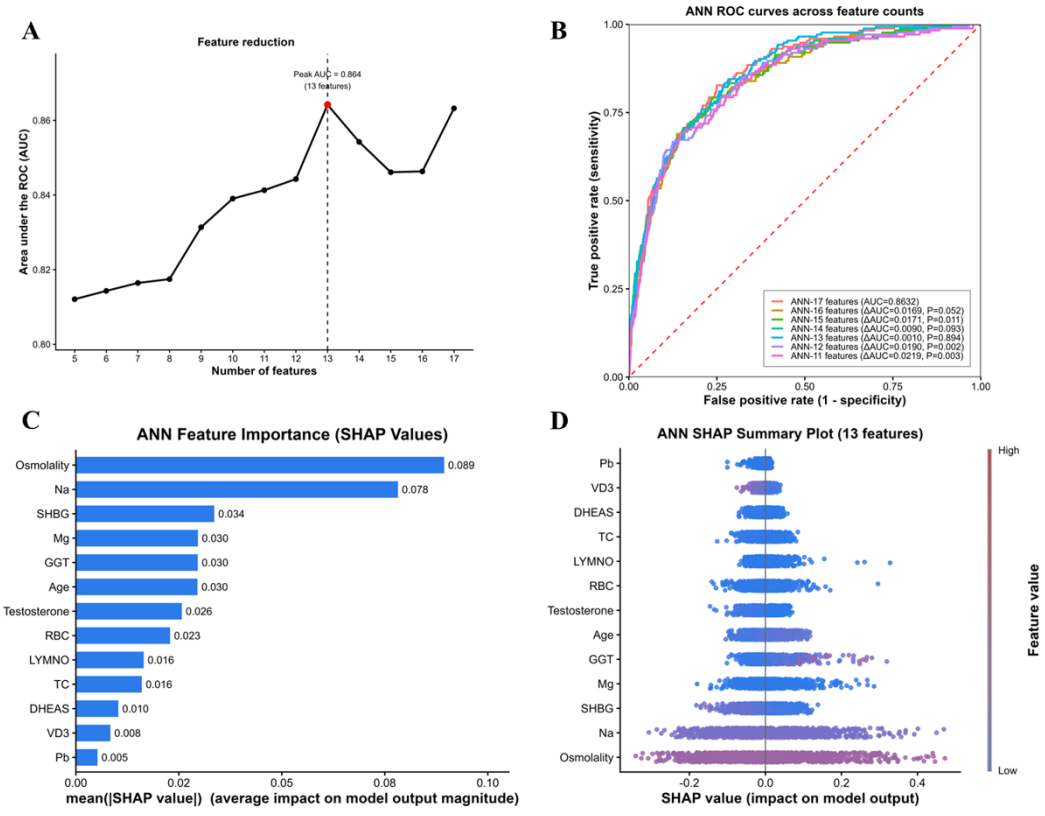

Figure 7: Feature selection and interpretability of the optimal ANN model

Note: (A) The feature reduction curve illustrating the area under the receiver operating characteristic curve (AUC) across different numbers of retained features. The peak AUC (0.864) was achieved with an optimal subset of 13 features. (B) Comparison of ROC curves for the ANN model utilizing different feature counts (ranging from 11 to 17). (C) Global feature importance ranking of the final 13-feature ANN model, evaluated by the mean absolute SHAP values. (D) SHAP summary plot demonstrating the distribution of SHAP values for each feature. Each dot represents a single patient, with the color indicating the actual feature value (red for high, blue for low) and the position on the x-axis indicating the impact on the model's output (risk of comorbidity).

3.8. Development and Web-Based Deployment of the Online Prediction Tool

To facilitate the clinical translation and practical application of the optimized ANN model, an interactive, web-based risk prediction calculator was developed, incorporating the 13 identified core predictors (Figure 8). The input interface integrates patient demographic and routine laboratory parameters, specifically osmolality, Na, Mg, SHBG, age, RBC, GGT, Pb, TC, LYMNO, VD3, testosterone, and DHEAS.

To ensure the standardization and accuracy of clinical data entry, the tool explicitly indicates the standard measurement units for each variable and features an "Input Patient Parameters" module for rapid verification by clinicians. Upon data entry, the system automatically calculates the individualized predicted probability of MAFLD and diabetes comorbidity. Furthermore, it allows users to specify a customized risk threshold (e.g., a default of 0.50) based on specific clinical requirements. Based on this threshold, the system provides a straightforward risk stratification warning (high vs. low risk). This web-based deployment successfully translates a complex machine learning algorithm into an accessible clinical decision support system (CDSS), substantially enhancing its practical utility in real-world healthcare settings.

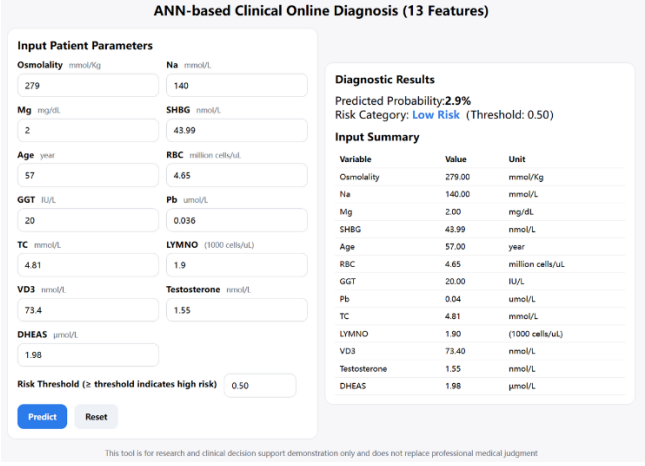

Figure 8: User interface of the web-based online diagnostic calculator developed from the optimized ANN model

Note: The left panel allows clinicians to input the specific values of the 13 selected predictive features. Upon clicking the “Predict” button, the right panel dynamically displays the calculated predicted probability of the comorbidity and stratifies the patient's risk category based on a customizable threshold (default = 0.50). An input summary is also generated for quick data verification. This interactive clinical decision support system translates the complex machine learning algorithm into an accessible point-of-care tool.

4. Discussion

This study developed and validated a machine learning-based comorbidity risk prediction model for MAFLD and DM using a nationally representative, multi-ethnic cohort from the NHANES 2021-2023 database. Using a dual feature selection strategy that combines LASSO regression and RF, 23 candidate variables were identified, and 10 machine learning algorithms were systematically compared on them. The ANN model demonstrated superior discriminative and calibration performance, achieving an AUC of 0.863, a Brier score of 0.073, and a non-significant Hosmer-Lemeshow test (P = 0.370) on the independent test set. Following SHAP-guided feature refinement, a parsimonious 13-feature ANN model was established without significant loss of discriminative performance (P = 0.894 vs. full-feature model), which was subsequently deployed as an open-access, web-based clinical decision support tool. These findings collectively demonstrate that integrating routine clinical and biochemical phenotypes within a nonlinear algorithmic framework can effectively quantify the early comorbidity risk of MAFLD and DM, offering a simple, non-invasive assessment tool to support the clinical implementation of the "dual metabolic-hepatic management" strategy and facilitate the early screening of high-risk populations.

Through SHAP explainability analysis, our study identified serum osmolality and Na levels as highly weighted predictors for the risk of comorbid MAFLD and DM. These findings suggest that for predicting the risk of MAFLD-DM comorbidity, systemic homeostatic dysregulation may offer superior value compared to classic liver enzymes like ALT and GGT, which demonstrated lower SHAP importance weights. This aligns perfectly with an emerging pathophysiological framework that repositions hyperosmolarity not merely as a downstream biochemical manifestation, but as an upstream driver of metabolic dysregulation. From a pathophysiological perspective, chronic hyperglycemia during prediabetes can induce a state of peripheral hyperosmolality. Rather than merely being a biochemical manifestation of metabolic dysregulation, this hyperosmotic environment may act as an upstream driver in the progression of comorbid MAFLD and DM by activating osmotic stress-related signaling pathways (such as the NFAT5/TonEBP axis, the NF-κB pathway, and oxidative stress cascades) [28][29][30]. Concurrently, even a mild elevation in serum osmolality potently triggers the hypothalamic release of arginine vasopressin (AVP) [31]. Upon specifically binding to V1a receptors on the hepatocyte surface, circulating AVP not only induces glycogenolysis and exacerbates systemic insulin resistance, but also upregulates key lipogenic enzymes to directly activate hepatic de novo lipogenesis (DNL) [32][33][34]. The activation of this neuroendocrine-metabolic axis constitutes one of the core mechanisms driving hepatocyte steatosis.

Regarding the role of serum sodium, its apparent 'protective effect' observed in the multivariable logistic regression should be interpreted with caution. Within the global feature space of non-linear models such as ANN, fluctuations in serum sodium are more likely to reflect the body's compensatory volume regulation in response to a chronic hyperosmolar environment, rather than a direct physiological protective mechanism. A

sustained hyperglycemic burden significantly enhances renal tubular reabsorption mediated by sodium-glucose cotransporters (SGLTs), inducing a compensatory redistribution of water and sodium across body compartments [35][36]. Subtle dynamic shifts in serum sodium essentially serve as a composite surrogate marker for the renal compensatory adaptation to osmotic stress. This complex, non-linear characteristic is one that traditional linear models struggle to effectively capture. In summary, the interplay between osmolality-driven AVP activation and the compensatory dysregulation of water and sodium orchestrates a pathogenic microenvironment that drives the mutual exacerbation of MAFLD and diabetes. This provides a systemic perspective, transcending isolated organ limitations, for understanding the “hepato-glycemic vicious cycle”.

The high enrichment of SHBG, testosterone, and DHEAS in the optimization model provides compelling evidence for the involvement of the 'liver-metabolic-endocrine axis' in the progression of comorbidities. Functioning as a critical hepatokine, SHBG is synthesized almost exclusively by hepatocytes, and its transcriptional activity is regulated by the liver's metabolic state [37][38]. During the pathogenesis and progression of MAFLD, ectopic lipid accumulation within hepatocytes, the local inflammatory microenvironment, and systemic hyperinsulinemia collectively exert potent transcriptional inhibitory signals on SHBG gene expression [39][40]. The decline in circulating SHBG levels likely reflects a liver-derived endocrine imbalance driven collectively by hepatocellular lipid accumulation, enhanced de novo lipogenesis (DNL), and exacerbated insulin resistance [41]. Therefore, low SHBG serves as an early indicator of the transition from simple steatosis to metabolic deterioration and heightened MASH risk in MAFLD. However, whether this reduction strictly predates conventional histological alterations warrants further confirmation through longitudinal cohorts, quantitative imaging, and histopathological validation. Furthermore, the inclusion of androgen-related indices such as DHEAS and androstenedione in the model suggests that the endogenous steroid hormone network may actively participate in the pathophysiological processes driving comorbidity progression, including ectopic lipid deposition, immune-inflammatory remodeling, and macrophage polarization. From a metabolic-endocrine perspective, this finding highlights the potential contribution of sex hormone imbalance to comorbidity progression. It also supports the need for future studies stratified by biological sex, age, and menopausal status to inform more individualized strategies for metabolic risk assessment and precision intervention. Compared to most comorbidity risk models that rely predominantly on traditional metabolic, inflammatory, or hepatorenal parameters, incorporating the steroid hormone dimension offers a novel, complementary perspective for both mechanistic interpretation and risk stratification.

The inclusion of GGT, magnesium, vitamin D, and selected trace elements further suggests that comorbidity development involves broader systemic processes, including hepatic oxidative stress, mineral metabolism, immune regulation, and environmental exposure [42][43][44][45][46]. These variables may complement traditional metabolic and endocrine markers, although their mechanistic contributions require further validation.

In clinical metabolism and hepatology, commonly used indices for assessing liver involvement, such as the FIB-4 index and NAFLD fibrosis score, are primarily designed to identify advanced fibrosis and may be insufficiently sensitive for detecting early-stage comorbid susceptibility. In addition, conventional logistic regression models may be limited when applied to highly correlated biochemical variables that are continuous in nature and biologically interconnected, as multicollinearity and nonlinear interactions can compromise model stability and discrimination [47]. In this context, the ANN-based model used in the present study offers methodological advantages by capturing complex nonlinear relationships and interactions among routine metabolic indicators. Compared with tree-based ensemble algorithms such as RF and ET, the ANN model maintained strong discriminative performance (AUC = 0.8632) and favorable calibration, as reflected by the Hosmer-Lemeshow test (P = 0.370). This balance between discrimination and calibration suggests that the model may provide reliable individualized risk estimation in heterogeneous real-world populations with chronic metabolic disorders.

Leveraging these methodological strengths, we successfully translated the optimized "ANN-13 feature model" into an interactive Clinical Decision Support System (CDSS) to bridge the long-standing disciplinary silos between diabetes and MAFLD management. In routine practice, the management of diabetes and MAFLD often remains fragmented across disciplines, leading to under-recognition of dysglycemia in patients with MAFLD or delayed evaluation of liver involvement in patients with diabetes. The proposed tool has potential clinical utility because all 13 included variables are derived from widely available routine blood biochemical tests, enabling accessible and non-invasive risk assessment without additional procedural burden or substantial extra cost. Through real-time individualized risk calculation, the web-based tool may assist clinicians in identifying high-risk individuals and prioritizing further evaluation, multidisciplinary management, and timely intervention. In particular, it may provide a practical basis for considering therapeutic strategies with potential dual metabolic and hepatic benefits, such as GLP-1 receptor agonists or SGLT2 inhibitors, in appropriate high-risk patients. Therefore, this ANN-based CDSS may serve as a scalable early-warning tool for improving risk stratification and supporting integrated management of diabetes-MAFLD comorbidity in real-world clinical settings.

Despite its strengths, including a relatively large sample size, a robust feature selection procedure, and comprehensive multidimensional model performance evaluation, this study has several limitations. First, the cross-sectional design of the NHANES database limits causal inference; therefore, the temporal sequence and causal relationships between the candidate features and the occurrence of MAFLD-DM comorbidity cannot be determined. Second, some features included in the model, particularly trace elements and sex hormone-related markers, may be influenced by dietary patterns, nutritional supplement use, potential medication history, renal

function status, and other unmeasured factors. Although we adjusted for known confounders as comprehensively as possible, residual confounding cannot be completely ruled out. Finally, model performance in this study was primarily evaluated using an internally split test set, and independent external validation is still lacking. Therefore, the generalizability of the model beyond the multiethnic North American population, particularly to East Asian Han populations and other groups with different genetic backgrounds, lifestyles, and disease spectra, requires further confirmation. Future studies should use multicenter, prospective longitudinal cohorts with dynamic follow-up data to externally validate, recalibrate, and assess the model's clinical utility, thereby facilitating its robust translation into real-world clinical practice.

5. Conclusions

In conclusion, based on a nationally representative cohort from the NHANES database, this cross-sectional study successfully developed and validated a highly robust ANN model to predict the concurrent risk of MAFLD and diabetes mellitus. By utilizing 13 accessible routine clinical and biochemical parameters, the optimized model demonstrated excellent discriminative and calibration performance. Importantly, SHAP interpretability analysis highlighted the critical predictive value of systemic homeostatic and endocrine dysregulation—particularly serum osmolality, sodium, and SHBG — offering novel pathophysiological insights that transcend traditional localized liver enzymes. To bridge the gap between complex algorithms and clinical application, we transformed this model into an open-access, web-based clinical decision support tool. This non-invasive and cost-effective calculator enables rapid individualized risk stratification, providing practical support for the implementation of the "integrated diabetes-liver co-management" strategy in real-world settings. Future prospective, multi-center longitudinal studies are warranted to externally validate this tool and further evaluate its generalizability across diverse populations.

CRediT authorship contribution statement

Rui Wang: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Yang Cheng**: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft. **Yun Wu**: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Fengqian Liu**: Investigation, Data curation. **Jing Yu**: Investigation, Data curation. **Jie Wang**: Formal analysis, Investigation, Data curation.

Declaration of generative AI and AI-assisted technologies in the writing process

The free versions of Grammarly were used for spell-checking and grammar corrections. After using these tools, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This study was supported by the Tianshan Innovation Team Project of Xinjiang Uygur Autonomous Region (Grant No. 2024D14017), the Distinguished Young Scientists Fund Project of Xinjiang Uygur Autonomous Region (Grant No. 2024D01E25), the National Natural Science Foundation of China (Grant No. 82260081), and the Graduate Research Innovation Project of Xinjiang Uygur Autonomous Region (Grant No. XJ2026G147).

Declaration of competing interests

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

Data Availability Statement

Data were available from the NHANES website via https://www.cdc.gov/nchs/nhanes/. All data generated or analyzed during this study are included in this article and its Supplementary Material files. Further enquiries can be directed to the corresponding author.

Acknowledgment

We acknowledge the efforts of the NHANES research team and thank all participants for their valuable time and involvement in the survey.

REFERENCES

[1] Le, M. H., Yeo, Y. H., Li, X., et al. (2022). 2019 Global NAFLD Prevalence: A Systematic Review and Meta-analysis. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 20(12), 2809-2817.e2828.

[2] Lim, G. E. H., Tang, A., Ng, C. H., et al. (2023). An Observational Data Meta-analysis on the Differences in Prevalence and Risk Factors Between MAFLD vs NAFLD. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 21(3), 619-629.e617.

[3] Miao, L., Targher, G., Byrne, C. D., et al. (2024). Current status and future trends of the global burden of MASLD. Trends in endocrinology and metabolism: TEM, 35(8), 697-707.

[4] Tang, A. S. P., Chan, K. E., Quek, J., et al. (2022). Non-alcoholic fatty liver disease increases risk of carotid atherosclerosis and ischemic stroke: An updated meta-analysis with 135,602 individuals. Clinical and molecular hepatology, 28(3), 483-496.

[5] Lin, H., Zhang, X., Li, G., et al. (2021). Epidemiology and Clinical Outcomes of Metabolic (Dysfunction)-associated Fatty Liver Disease. Journal of clinical and translational hepatology, 9(6), 972-982.

[6] Thomas, J. A., Kendall, B. J., El-Serag, H. B., et al. (2024). Hepatocellular and extrahepatic cancer risk in people with non-alcoholic fatty liver disease. The lancet. Gastroenterology & hepatology, 9(2), 159-169.

[7] Sun, H., Saeedi, P., Karuranga, S., et al. (2022). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for

2021 and projections for 2045. Diabetes research and clinical practice, 183, 109119.

[8] Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. (2024). Lancet (London, England), 404(10467), 2077-2093.

[9] Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. (2023). Lancet (London, England), 402(10397), 203-234.

[10] Yang, T., Qi, F., Guo, F., et al. (2024). An update on chronic complications of diabetes mellitus: from molecular mechanisms to therapeutic strategies with a focus on metabolic memory. Molecular medicine (Cambridge, Mass.), 30(1), 71.

[11] Adams, L. A., Anstee, Q. M., Tilg, H., et al. (2017). Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. Gut, 66(6), 1138-1153.

[12] Leite, N. C., Villela-Nogueira, C. A., Cardoso, C. R., et al. (2014). Non-alcoholic fatty liver disease and diabetes: from physiopathological interplay to diagnosis and treatment. World journal of gastroenterology, 20(26), 8377-8392.

[13] Li, Y., Yang, P., Ye, J., et al. (2024). Updated mechanisms of MASLD pathogenesis. Lipids in health and disease, 23(1), 117.

[14] Moretti, V., Romeo, S., & Valenti, L. (2024). The contribution of genetics and epigenetics to MAFLD susceptibility. Hepatology international, 18(Suppl 2), 848-860.

[15] Lee, C. H., Lui, D. T., & Lam, K. S. (2022). Non-alcoholic fatty liver disease and type 2 diabetes: An update. Journal of diabetes investigation, 13(6), 930-940.

[16] Huang, D. Q., Noureddin, N., Ajmera, V., et al. (2023). Type 2 diabetes, hepatic decompensation, and hepatocellular carcinoma in patients with non-alcoholic fatty liver disease: an individual participant-level data meta-analysis. The lancet. Gastroenterology & hepatology, 8(9), 829-836.

[17] Mantovani, A., Csermely, A., Petracca, G., et al. (2021). Non-alcoholic fatty liver disease and risk of fatal and non-fatal cardiovascular events: an updated systematic review and meta-analysis. The lancet. Gastroenterology & hepatology, 6(11), 903-913.

[18] Kapoor, N., & Kalra, S. (2023). Metabolic-Associated Fatty Liver Disease and Diabetes: A Double Whammy. Endocrinology and metabolism clinics of North America, 52(3), 469-484.

[19] Cusi, K., Abdelmalek, M. F., Apovian, C. M., et al. (2025). Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) in People With Diabetes: The Need for Screening and Early Intervention. A Consensus Report of the American Diabetes Association. Diabetes care, 48(7), 1057-1082.

[20] Cusi, K., Isaacs, S., Barb, D., et al. (2022). American Association of Clinical Endocrinology Clinical Practice Guideline for the Diagnosis and Management of Nonalcoholic Fatty Liver Disease in Primary Care and Endocrinology Clinical Settings: Co-Sponsored by the American Association for the Study of Liver Diseases (AASLD). Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists, 28(5), 528-562.

[21] Cho, H., She, J., De Marchi, D., et al. (2024). Machine Learning and Health Science Research: Tutorial. Journal of medical Internet research, 26, e50890.

[22] Anthonimuthu, D. J., Hejlesen, O., Zwisler, A. O., et al. (2024). Application of Machine Learning in Multimorbidity Research: Protocol for a Scoping Review. JMIR research protocols, 13, e53761.

[23] Patel, C. J., Pho, N., McDuffie, M., et al. (2016). A database of human exposomes and phenomes from the US National Health and Nutrition Examination Survey. Scientific data, 3, 160096.

[24] Siddiqui, M. S., Vuppalanchi, R., Van Natta, M. L., et al. (2019). Vibration-Controlled Transient Elastography to Assess Fibrosis and Steatosis in Patients With Nonalcoholic Fatty Liver Disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 17(1), 156-163.e152.

[25] Eslam, M., Newsome, P. N., Sarin, S. K., et al. (2020). A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. Journal of hepatology, 73(1), 202-209.

[26] Eslam, M., Fan, J. G., Yu, M. L., et al. (2025). The Asian Pacific association for the study of the liver clinical practice guidelines for the diagnosis and management of metabolic dysfunction-associated fatty liver disease. Hepatology international, 19(2), 261-301.

[27] Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024. (2024). Diabetes care, 47(Suppl 1), S20-s42.

[28] Peng, L., & Ai, M. (2025). Association of serum osmolality with metabolic dysfunction-associated steatotic liver disease in adults. BMC gastroenterology, 25(1), 693.

[29] Lee, J. H., Song, H., Yoo, E. J., et al. (2026). Hepatocyte TonEBP promotes metabolic stress-induced hepatic fibroinflammation involving transcriptional activation of ELR^+ CXC chemokines. Cell death discovery, 12(1).

[30] Lee, H. H., Jeong, G. W., Ye, B. J., et al. (2022). TonEBP in Myeloid Cells Promotes Obesity-Induced Insulin Resistance and Inflammation Through Adipose Tissue Remodeling. Diabetes, 71(12), 2557-2571.

[31] Kim, D. H., Kim, K. K., Lee, T. H., et al. (2021). Transcription Factor TonEBP Stimulates Hyperosmolality-Dependent Arginine Vasopressin Gene Expression in the Mouse Hypothalamus. Frontiers in endocrinology, 12, 627343.

[32] Szczepanska-Sadowska, E., Cudnoch-Jędrzejewska, A., & Żera, T. (2024). Molecular Interaction Between Vasopressin and Insulin in

Regulation of Metabolism: Impact on Cardiovascular and Metabolic Diseases. International journal of molecular sciences, 25(24).

[33] Student, J., Sowers, J., & Lockette, W. (2022). THIRSTY FOR FRUCTOSE: Arginine Vasopressin, Fructose, and the Pathogenesis of Metabolic and Renal Disease. Frontiers in cardiovascular medicine, 9, 883365.

[34] Andres-Hernando, A., Jensen, T. J., Kuwabara, M., et al. (2021). Vasopressin mediates fructose-induced metabolic syndrome by activating the V1b receptor. JCI insight, 6(1).

[35] Kim, N. H., & Kim, N. H. (2022). Renoprotective Mechanism of Sodium-Glucose Cotransporter 2 Inhibitors: Focusing on Renal Hemodynamics. Diabetes & metabolism journal, 46(4), 543-551.

[36] Umpierrez, G. E., Davis, G. M., ElSayed, N. A., et al. (2024). Hyperglycemic Crises in Adults With Diabetes: A Consensus Report. Diabetes care, 47(8), 1257-1275.

[37] Simons, P., Valkenburg, O., Stehouwer, C. D. A., et al. (2021). Sex hormone-binding globulin: biomarker and hepatokine? Trends in endocrinology and metabolism: TEM, 32(8), 544-553.

[38] Bourebaba, N., Ngo, T., Śmieszek, A., et al. (2022). Sex hormone binding globulin as a potential drug candidate for liver-related metabolic disorders treatment. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 153, 113261.

[39] Stefan, N., Schick, F., Birkenfeld, A. L., et al. (2023). The role of hepatokines in NAFLD. Cell metabolism, 35(2), 236-252.

[40] Simons, P., Valkenburg, O., van de Waarenburg, M. P. H., et al. (2023). Serum sex hormone-binding globulin is a mediator of the association between intrahepatic lipid content and type 2 diabetes: the Maastricht Study. Diabetologia, 66(1), 213-222.

[41] Brianso-Llort, L., Saéz-Lopez, C., Alvarez-Guaita, A., et al. (2024). Recent Advances on Sex Hormone-Binding Globulin Regulation by Nutritional Factors: Clinical Implications. Molecular nutrition & food research, 68(14), e2400020.

[42] Xing, Y., Chen, J., Liu, J., et al. (2022). Associations Between GGT/HDL and MAFLD: A Cross-Sectional Study. Diabetes, metabolic syndrome and obesity : targets and therapy, 15, 383-394.

[43] Wu, W., Gong, M., Liu, P., et al. (2025). Hypomagnesemia: exploring its multifaceted health impacts and associations with blood pressure regulation and metabolic syndrome. Diabetology & metabolic syndrome, 17(1), 217.

[44] Adetunji, A. G., & Obeng-Gyasi, E. (2024). Investigating the Interplay of Toxic Metals and Essential Elements in Liver Disease. International journal of environmental research and public health, 21(6).

[45] Fenercioglu, A. K. (2024). The Anti-Inflammatory Roles of Vitamin D for Improving Human Health. Current issues in molecular biology, 46(12), 13514-13525.

[46] Piuri, G., Zocchi, M., Della Porta, M., et al. (2021). Magnesium in Obesity, Metabolic Syndrome, and Type 2 Diabetes. Nutrients, 13(2).

[47] Ma, J., Dhiman, P., Qi, C., et al. (2023). Poor handling of continuous predictors in clinical prediction models using logistic regression: a systematic review. Journal of clinical epidemiology, 161, 140-151.