

RESEARCH

Open Access



Development and validation of machine learning models for predicting low muscle mass in patients with obesity and diabetes

Jiaying Ge¹, Siqi Sun¹, Jiangping Zeng², Yujie Jing¹, Huihui Ma³, Chunhua Qian¹, Ran Cui¹, Shen Qu¹ and Hui Sheng^{1*}

Abstract

Background and aims Low muscle mass (LMM) is a critical complication in patients with obesity and diabetes, exacerbating metabolic and cardiovascular risks. Novel obesity indices, such as the body roundness index (BRI), conicity index, and relative fat mass, have shown promise for assessing body composition. This study aimed to investigate the associations of these indices with LMM and to develop machine learning models for accurate and accessible LMM prediction.

Method Data from NHANES 2011–2018 ($n = 2,176$) were analyzed. Obesity was defined by body fat percentage, and LMM was determined using skeletal muscle mass index thresholds adjusted for BMI. Predictive models were developed using logistic regression, random forest, and other algorithms, with feature selection via LASSO regression. Validation included NHANES 2005–2006 data ($n = 310$). Model performance was evaluated using AUROC, Brier scores, and SHapley Additive exPlanations (SHAP) for feature importance.

Results BRI was independently associated with LMM (odds ratio 1.39, 95% confidence interval 1.22–1.58; $P < 0.001$). Eight features were included in the random forest model, which achieved excellent discrimination (AUROC = 0.721 in the validation set) and calibration (Brier score = 0.184). Feature importance analysis highlighted BRI, creatinine, race, age, and HbA1c as key contributors to the model's predictive performance. SHAP analysis emphasized BRI's role in predicting LMM. An online prediction tool was developed.

Conclusions BRI is a significant predictor of LMM in patients with obesity and diabetes. The random forest model demonstrated strong performance and offers a practical tool for early LMM detection, supporting clinical decision-making and personalized interventions.

Keywords Low muscle mass, Diabetes, Obesity, NHANES, Machine learning

*Correspondence:

Hui Sheng
shenghui@tongji.edu.cn

¹Department of Endocrinology and Metabolism, Shanghai Tenth People's Hospital, School of Medicine, Tongji University, Shanghai 200072, China

²Department of Endocrinology and Metabolism, Suzhou Dushu Lake Hospital, The Fourth Affiliated Hospital of Soochow University, Medical Center of Soochow University, Suzhou 215123, Jiangsu, China

³Department of Endocrinology, Jinan Central Hospital Affiliated to Shandong First Medical University, Jinan 250000, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

The global incidence of diabetes continues to rise, projected to increase from 415 million people in 2015 to 642 million by 2040 [1]. Diabetes is primarily caused by insulin deficiency and resistance, which lead to significant metabolic disturbances [2]. Sarcopenia—an age-related decline in muscle mass and function—progressed at about 8% per decade after age 40 and 15–25% after age 70 [3, 4], affecting 7–29% of patients with diabetes [5, 6, 7]. Coexisting obesity further accelerates muscle and metabolic deterioration, elevating risks of disability, cardiovascular disease, and mortality. Moreover, the decline in muscle mass worsens metabolic dysregulation in patients with diabetes [8, 9, 10, 11]. Sarcopenic obesity (increased fat with reduced muscle mass) aggravates hyperglycemia and insulin resistance, intensifying adverse health outcomes [12, 13, 14]. Although imaging techniques like dual-energy X-ray absorptiometry (DXA) and magnetic resonance imaging (MRI) accurately identify low muscle mass (LMM), their high cost and technical demands limit widespread application. Our previous studies have used biochemical markers such as the creatinine/cystatin C ratio, sarcopenia index, and soluble interleukin-2 receptor to detect LMM or sarcopenia [15, 16], yet clinically feasible tools remain scarce in individuals with obesity and diabetes.

Recently, several novel obesity indices, including the body roundness index (BRI), conicity index (C-index), and relative fat mass (RFM), have emerged as enhanced measures of fat distribution and body composition [17, 18, 19]. Thomas [20] et al. proposed BRI as an anthropometric index associated with obesity, and the C-index provided insights into fat distribution, particularly abdominal fat accumulation [21]. RFM, which estimates total body fat percentage more accurately than body mass index (BMI), offers a better representation of fat distribution [17]. Although these indices are closely linked to chronic diseases such as diabetes and cardiovascular disease [22, 23, 24], their usefulness in detecting low LMM among patients with obesity and diabetes has not been thoroughly investigated.

Therefore, this study aimed to explore the association between novel obesity indices and LMM in this high-risk population. By leveraging multiple machine learning models, this study aimed to develop a practical predictive tool for early LMM detection, ultimately enhancing clinical management and personalized treatment strategies for individuals with obesity and diabetes.

Materials and methods

Study population and design

Data for this analysis were drawn from the 2005–2006 and 2011–2018 National Health and Nutrition Examination Survey (NHANES) datasets. The NHANES initiative

aims to examine the health and nutrition of adults and children in the United States through interviews and physical examinations. The survey includes demographic, dietary, socioeconomic, and health-related questions. Approval for this study was provided by the Research Ethics Review Board of the National Center for Health Statistics and all participants provided written informed consent. Additional details are available at <https://www.cdc.gov/nchs/nhanes/index.htm>.

This study was based on NHANES patient data between 2011 and 2018 ($n=78188$). Patients with diabetes were included in the study if they met one of the following criteria: Hemoglobin A1c (HbA1c) $\geq 6.5\%$, fasting glucose ≥ 7 mmol/L, or self-reported use of antidiabetic medication based on a questionnaire [25]. Obesity was defined as a total body fat percentage of $\geq 25\%$ for men and $\geq 30\%$ for women [26]. The following exclusion criteria were applied to identify the final participants: (1) age < 40 or > 70 years; (2) not meeting the diagnostic criteria for diabetes and obesity; (3) missing DXA data; (4) missing waist circumference (WC), weight, or height data; (5) missing sex data; and (6) missing important covariate data, including drinking status, smoking status, hypertension, and medication use. External validation was performed using NHANES data from 2005 to 2006 ($n=10348$). The same inclusion and exclusion criteria were applied to this cohort, resulting in 310 participants in the external validation dataset. A flow diagram illustrating the inclusion process for both the development and external validation cohorts is displayed in Figure S1.

Measurements and definition of LMM

NHANES used a Hologic QDR-4500 fan-beam densitometer (Hologic Inc., Bedford, MA, USA) to measure body composition using DXA. Appendicular skeletal muscle mass (ASM) was calculated as the total of the muscle mass in both the arms and legs. The Foundation for the National Institutes of Health recommends that ASM be adjusted for BMI when the skeletal muscle mass index (SMI) is used as a diagnostic metric for LMM. Men with an SMI < 0.789 and women with an SMI < 0.512 were classified as having LMM [27].

Variable estimation

Data on age, sex, race, height, weight, WC, BMI, blood pressure, smoking and drinking status, complete blood count, biochemical information, and medication use were obtained from the publicly available NHANES demographic and release datasets. Race was classified into five categories: Mexican American, other Hispanic, non-Hispanic white, non-Hispanic black, and other. BMI was calculated as weight divided by height squared (kg/m^2). Hypertension was defined as the average of two blood pressure readings released by the NHANES, with

a mean systolic pressure > 140 mmHg or diastolic pressure > 90 mmHg.

BRI, RFM, and C-index

$$\text{BRI} = 364.2 - 365.5 \times \sqrt{\frac{1 - \left(\frac{\text{WC}(\text{m})}{2\pi}\right)^2}{(0.5 \times \text{height}(\text{m}))^2}}$$

$$\text{RFM (men)} = 64 - 20 \times \frac{\text{height}(\text{m})}{\text{waist}(\text{m})}; \quad \text{RFM (women)} = 76 - 20 \times \frac{\text{height}(\text{m})}{\text{waist}(\text{m})}; \quad \text{C-index} = 0.109^{-1} \text{WC}(\text{m}) \frac{\text{weight}(\text{kg})^{-1/2}}{\text{height}(\text{m})}.$$

Statistical analyses

Data normality was assessed using the Kolmogorov–Smirnov test and Q-Q. Continuous variables with a normal distribution are presented as means ± standard deviations, whereas non-normally distributed variables are expressed as medians (interquartile ranges). Categorical variables are summarized as frequencies and percentages. Depending on distribution, continuous variables were compared with the independent samples t-test or Mann–Whitney U test. The chi-squared test or Fisher’s exact test was used to assess differences in categorical variables. Weighted logistic regression was conducted to identify independent risk factors, with sample weights adjusted to the 4-year survey cycle (WTMEC4YR). All statistical analyses were performed in R software (version 4.2.2), and statistical significance was set at $P < 0.050$ (two-tailed).

Participants was randomly divided into training (70%) and validation (30%) subsets using the R function `sample`, with a fixed random seed (123) to ensure reproducibility. To address sample imbalance, the synthetic minority oversampling technique algorithm (SMOTE) was applied. SMOTE was applied exclusively to the training dataset. Model performance was subsequently evaluated through 10-fold cross-validation to confirm that the synthetic data did not lead to systematic bias or inflated performance estimates. Missing data were imputed using the k-nearest neighbors approach (Table S1). To ensure robustness, we performed sensitivity analyses comparing the means of numerical variables before and after imputation, calculating the percentage difference between observed and imputed values (Table S2). Variables with $P < 0.100$ in univariate analysis were retained as candidates for multivariate analysis. Feature selection was carried out using the least absolute shrinkage and selection operator (LASSO). To identify the optimal penalty parameter (λ), we employed 10-fold cross-validation using the `cv.glmnet` function (glmnet package). The one standard error rule ($\lambda.1se$) was then applied to balance model complexity and predictive performance, with the selected $\lambda.1se$ subsequently used for feature selection.

Multicollinearity was assessed via the variance inflation factor (VIF). Six machine learning methods—logistic regression, decision tree, support vector machine, random forest, categorical boosting (CatBoost), and extreme gradient boosting (XGBoost)—were used to develop predictive models. Performance was evaluated using accuracy, precision, F1 score, recall, and area under the receiver operating characteristic (AUROC) curve. Calibration curves were generated to assess the calibration performance of the predictive models and compare the relationship between the actual values and model predictions. Decision curve analysis and Brier scores were used to evaluate the model fit.

Results

Characteristics of participants according to LMM status

The baseline characteristics of the final participants ($n = 2,176$) are presented in Table 1, with 21.3% of them having LMM. Weighted results showed that the median age of participants with LMM was 52.58 years (45.1% men); the median age of the participants with normal muscle mass (NMM) was 53 years (40.2% men). The median BRI and RFM values for the LMM and NMM groups were 9.11 vs. 6.73 and 45.01 vs. 38.24, respectively, with significant differences ($P < 0.010$); the C-index medians were 1.01 and 1.08, with no significant difference ($P = 0.086$). Additionally, significant differences were observed between the LMM and NMM groups in several baseline characteristics, including BMI, white blood cell count, globulin, creatinine, HbA1c, SMI, total body fat percentage, race, and drinking status ($P < 0.050$).

Correlation between BRI, RFM, C-index, and LMM

We used logistic regression analysis to explore the association between BRI, RFM, C-index, and LMM. The ORs and 95% CIs were calculated to quantify this association (Fig. 1). BRI showed the strongest positive association with LMM (odds ratio [OR] 1.39, 95% confidence interval [CI]: 1.22–1.58, $P < 0.001$). RFM also showed an association with LMM, although weaker (OR 1.07, 95% CI: 1.02–1.11, $P = 0.003$). However, the correlation between the C-index and LMM was not significant (OR 0.33, 95% CI: 0.07–1.52, $P = 0.151$). Given the potential multicollinearity between the BRI, RFM, and C-index, the BRI, which had the strongest association with LMM, will be included in subsequent studies to clarify its role as a predictive variable for LMM.

Development and validation of predictive models

We performed a weighted univariate regression analysis to identify independent factors significantly associated with LMM in patients with obesity and diabetes (Table 2). The results showed that higher WC, BMI, BRI, and alcohol consumption were independent risk factors

Table 1 Weighted characteristics of participants in study

Characteristic	Normal muscle mass (n = 1713)	Low muscle mass (n = 463)	P-value
Age (years)	53.00 (48.00, 56.00)	52.58 (44.00, 55.00)	0.420
WC (cm)	113.00 (102.98, 124.30)	117.80 (108.50, 135.50)	0.067
BMI (kg/m ²)	33.70 (29.20, 37.30)	37.80 (32.30, 43.70)	<0.001***
White blood cell count (10 ⁹ /L)	7.90 (6.70, 9.10)	8.30 (7.40, 9.40)	0.045*
Lymphocyte count (10 ⁹ /L)	2.20 (1.70, 2.70)	2.20 (2.00, 2.70)	0.146
Monocyte count (10 ⁹ /L)	0.60 (0.50, 0.70)	0.60 (0.50, 0.70)	0.551
Neutrophil count (10 ⁹ /L)	4.60 (3.70, 5.90)	5.05 (4.30, 6.00)	0.090
Hemoglobin (g/L)	14.40 (13.30, 15.40)	14.30 (13.20, 15.00)	0.473
Red cell distribution width (%)	13.50 (12.90, 14.10)	13.60 (13.20, 14.10)	0.411
Fasting glucose (mmol/L)	8.99 (7.11, 12.56)	7.12 (6.27, 9.63)	0.097
Albumin (g/L)	42.00 (40.00, 44.00)	41.00 (40.00, 44.00)	0.898
Globulin (g/L)	28.00 (26.00, 31.00)	29.00 (27.00, 33.00)	0.039*
Creatinine (umol/L)	74.26 (63.65, 90.17)	70.72 (61.88, 76.02)	0.005**
Uric acid (umol/L)	321.20 (255.80, 392.60)	327.10 (255.80, 379.58)	0.851
HbA1c (%)	7.20 (6.30, 8.90)	6.60 (6.10, 8.60)	0.017*
ALT (U/L)	25.00 (19.00, 37.00)	29.00 (23.00, 43.00)	0.070
AST (U/L)	25.00 (19.00, 30.00)	24.00 (20.00, 34.00)	0.749
ALP (U/L)	70.00 (57.00, 89.00)	72.00 (57.00, 91.00)	0.583
Low-density lipoprotein (mmol/L)	2.48 (1.94, 3.38)	2.87 (2.28, 4.03)	0.172
High-density lipoprotein (mmol/L)	1.14 (0.96, 1.29)	1.11 (0.98, 1.34)	0.628
Total cholesterol (mmol/L)	4.68 (3.88, 5.59)	4.60 (4.14, 5.87)	0.306
Triglyceride (mmol/L)	2.16 (1.21, 3.16)	1.67 (1.46, 2.36)	0.279
Systolic pressure (mmHg)	123.00 (113.00, 132.00)	123.00 (116.00, 138.00)	0.259
Diastolic pressure (mmHg)	74.00 (68.00, 80.00)	71.00 (67.00, 77.00)	0.198
SMI	0.80 (0.61, 0.92)	0.51 (0.49, 0.74)	<0.001***
Total body fat percent (%)	35.50 (30.50, 41.50)	41.80 (35.90, 46.20)	<0.001***
BRI	6.73 (5.43, 8.85)	9.11 (7.60, 10.99)	<0.001***
RFM	38.24 (32.13, 47.48)	45.01 (37.13, 49.41)	0.001**
C-index	1.08 (0.94, 1.25)	1.01 (0.81, 1.20)	0.086
Gender, n (%)			0.395
Female	940 (54.9)	277 (59.8)	
Male	773 (45.1)	186 (40.2)	
Race, n (%)			0.005**
Mexican American	227 (13.3)	142 (30.7)	
Other Hispanic	168 (9.8)	80 (17.3)	
Non-Hispanic White	574 (33.5)	161 (34.8)	
Non-Hispanic Black	540 (31.5)	47 (10.2)	
Other race	204 (11.9)	33 (7.1)	
Drinking status, n (%)			0.046*
Drinker	629 (36.7)	123 (26.6)	
Non-drinker	1084 (63.3)	340 (73.4)	
Smoking status, n (%)			0.860
Smoker	917 (53.5)	252 (54.4)	
Non-smoker	796 (46.5)	211 (45.6)	
Use of insulin, n (%)			0.275
Yes	1185 (69.2)	348 (75.2)	
No	528 (30.8)	115 (24.8)	
Use of metformin, n (%)			0.882
Yes	1418 (87.4)	386 (87.1)	
No	204 (12.6)	57 (12.9)	

Data are presented as median (IQR), or n (%). Abbreviations: WC, Waist circumference; BMI, Body mass index; HbA1c, glycated hemoglobin; ALT, Alanine transaminase; AST, Alanine aminotransferase; ALP, Alkaline phosphatase; SMI, skeletal muscle mass index; BRI, body roundness index; RFM, relative fat mass; C-index, conicity index; IQR, interquartile range. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

Forest Plot of BRI, RFM, and C-index with LMM (Weighted)

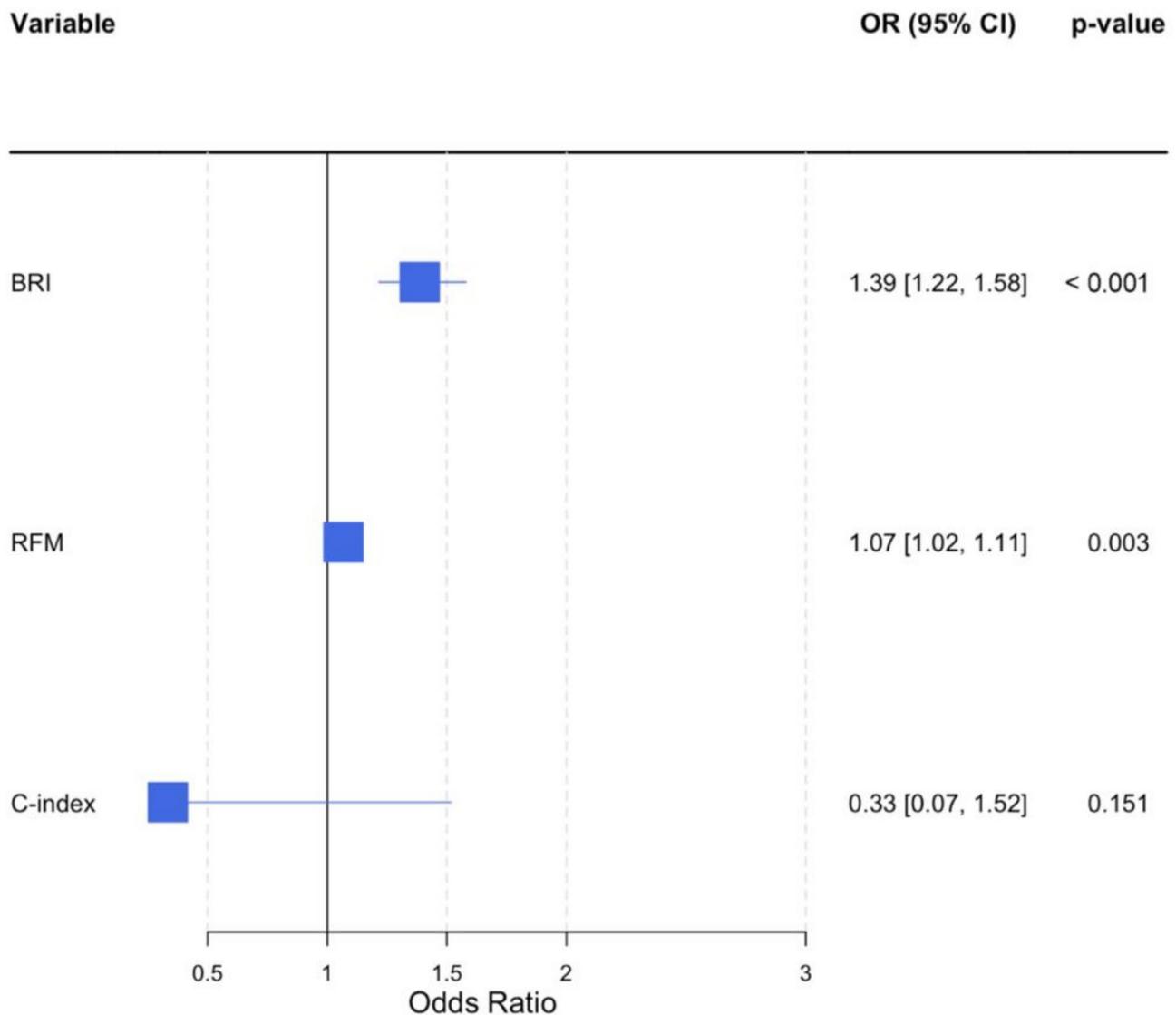


Fig. 1 Forest plot of BRI, RFM, and C-index with LMM. Abbreviations: OR, odds ratio; CI, confidence intervals; BRI, body roundness index; RFM, relative fat mass; C-index, conicity index; LMM, low muscle mass

for LMM ($P < 0.050$). Additionally, non-Hispanic white, non-Hispanic black and other ethnic groups were negatively associated with LMM, suggesting that these ethnic groups may have protective effects ($P < 0.050$). In terms of biochemical markers, elevated creatinine and HbA1c levels were independent protective factors against LMM ($P < 0.050$). Higher globulin levels were positively associated with LMM; however, this relationship was not significant.

We further analyzed the independent factors for LMM identified in the univariate regression analysis ($P < 0.100$) [28]. BMI and WC were removed to avoid potential

multicollinearity with the BRI. Additionally, based on clinical relevance and existing literature [29], age and sex were included in the model as covariates. We used the least absolute shrinkage and selection operator (LASSO) analysis to further filter the feature variables for LMM. The optimal penalty parameter (λ) was identified using the glmnet package. Variables with non-zero coefficients under the λ were retained. For categorical variables, dummy variables were encoded to ensure that different categories were accurately represented in the LASSO analysis. To minimize redundancy, we merge multiple dummy variables within the same category. Variables

Table 2 Weighted univariate analysis of variables with LMM

Variable	OR	95%CI	Pvalue
Age (years)	0.965	0.897–1.039	0.335
Waist circumference (cm)	1.024	1.002–1.046	0.037*
BMI (kg/m ²)	1.092	1.044–1.143	<0.001***
White blood cell count (10 ⁹ /L)	1.072	0.963–1.194	0.196
Lymphocyte count (10 ⁹ /L)	1.286	0.832–1.986	0.251
Monocyte count (10 ⁹ /L)	1.787	0.309–10.341	0.509
Neutrophil count (10 ⁹ /L)	1.058	0.938–1.194	0.350
Hemoglobin (g/L)	0.914	0.717–1.165	0.460
Red blood cell distribution width (%)	1.004	0.786–1.282	0.977
Fasting glucose (mmol/L)	0.892	0.771–1.032	0.122
Albumin (g/L)	1.017	0.932–1.109	0.698
Globulin (g/L)	1.057	0.994–1.123	0.075
Creatinine (umol/L)	0.977	0.962–0.992	<0.01**
Uric acid (umol/L)	1.000	0.997–1.004	0.829
HbA1c (%)	0.841	0.726–0.974	0.022*
ALT (U/L)	1.003	0.994–1.013	0.467
AST (U/L)	0.999	0.994–1.004	0.742
ALP (U/L)	1.005	0.994–1.016	0.370
Low-density lipoprotein (mmol/L)	1.376	0.869–2.178	0.169
High-density lipoprotein (mmol/L)	1.167	0.471–2.889	0.733
Total cholesterol (mmol/L)	1.025	0.927–1.133	0.627
Triglyceride (mmol/L)	0.895	0.780–1.028	0.115
Systolic pressure (mmHg)	1.009	0.995–1.022	0.194
Diastolic pressure (mmHg)	0.984	0.959–1.010	0.231
BRI	1.387	1.217–1.581	<0.001***
Gender (female)	1.393	0.639–3.033	0.396
Race			<0.010**
Other Hispanic	0.620	0.222–1.731	0.352
Non-Hispanic White	0.381	0.160–0.904	0.030*
Non-Hispanic Black	0.124	0.043–0.359	<0.001***
Other Race	0.204	0.064–0.651	<0.01**
Drinking status (drinker)	1.778	1.006–3.144	0.048*
Smoking status (smoker)	0.932	0.421–2.064	0.860
Use of insulin (yes)	0.657	0.304–1.419	0.277
Use of metformin (yes)	1.023	0.754–1.388	0.882

Abbreviations: LMM, low muscle mass; WC, waist circumference; BMI, body mass index; HbA1c, glycated hemoglobin; ALT, alanine transaminase; AST, alanine aminotransferase; ALP, alkaline phosphatase; SMI, skeletal muscle mass index; BRI, body roundness index; RFM, relative fat mass; C-index, conicity index; IQR, interquartile range. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

with the variance inflation factor (VIF) > 5 were considered to have significant multicollinearity. Eight nonzero feature variables were selected to construct a predictive model for LMM risk (Figure S2).

The dataset was randomly divided into training and validation sets at a 7:3 ratio. The baseline characteristics of the two sets are presented in Table S3, which demonstrates no significant differences in key variables. This ensures the representativeness and generalizability of the model's performance across datasets.

To ensure optimal performance for each machine learning model, we used Bayesian optimization to tune the key model parameters and conducted 10 iterations to

identify the optimal parameter set (Table S4). Based on this, the final model was trained. The model performance was evaluated through 10-fold cross-validation on the internal training set. As shown in Fig. 2, the random forest demonstrated the highest clinical efficiency, achieving an area under the receiver operating characteristic (AUROC) of 0.994 (95% CI 0.992–0.995), while categorical boosting (CatBoost) followed closely with an AUROC of 0.985 (95% CI 0.975–0.996).

A further evaluation of the stability and generalization ability of the six predictive models was conducted for the internal and external validation sets. With an AUROC of 0.980 (95% CI 0.970–0.990), the random forest model demonstrated the best clinical predictive performance in the internal validation set (Fig. 3). Table 3 illustrates that the random forest model achieved high accuracy, precision, recall, and F1 score in both validation sets. The model also exhibited superior calibration performance, with a Brier score of 0.071 for the internal set and 0.184 for the external set (Figure S3, Figure S4). We further analyzed the net benefits of each model under different decision scenarios (Figure S5). The random forest model displayed higher net benefits across a broad range of thresholds, particularly in the 10–40% threshold probability range. Random forest showed the highest clinical predictive utility; therefore, it was selected as the optimal model. The SHapley Additive exPlanations (SHAP) feature importance for the random forest model is illustrated in Fig. 4, where the features are ranked from highest to lowest based on their mean absolute SHAP values. The five most important features were BRI, age, race, creatinine level, and HbA1c.

Risk category

Using the optimal threshold (0.305) derived from the random forest model via Youden's index, patients were stratified into high-risk (> 0.305) and low-risk (≤ 0.305) groups (Table 4). In the low-risk group, the predicted risk values (14.38%) were closely aligned with the observed risk values (14.04%), demonstrating the model's high predictive accuracy for this group. In the high-risk group, the predicted risk values (54.89%) were moderately higher than the observed risk values (43.17%), with a difference of 11.72%. Despite this discrepancy, the model maintained good discriminatory power in distinguishing between the two groups. The observed difference in LMM risk between the high-risk and low-risk groups was statistically significant ($P < 0.001$). Subsequently, an online predictive calculator was developed to facilitate the clinical application of LMM diagnosis: <https://sarcopeniadiagnosis.shinyapps.io/LowMuscleMassDiagnosis/>.

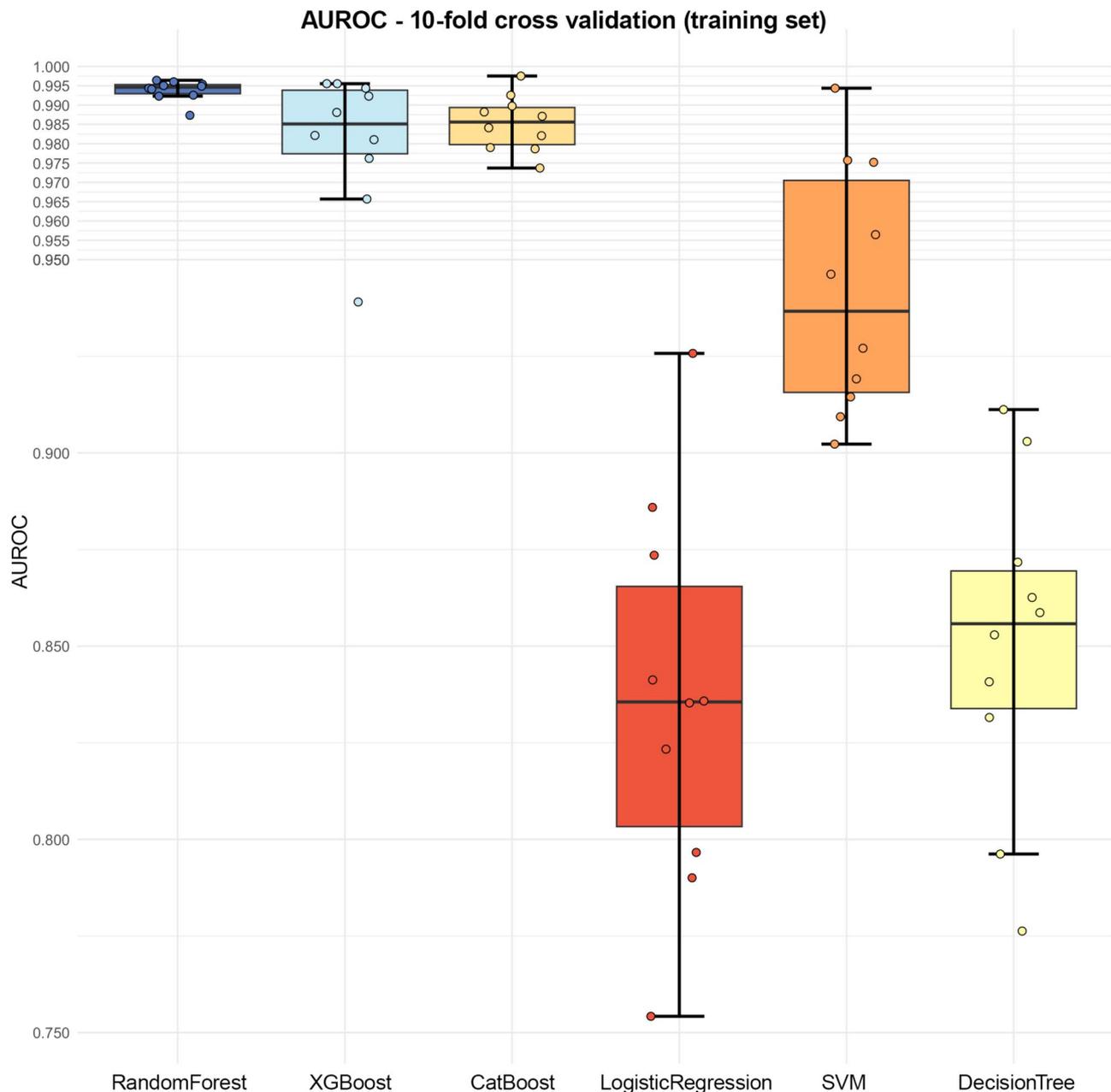


Fig. 2 Clinical diagnostic value of 6 machine learning models (10-fold cross-validation) in the training set. Abbreviations: AUROC, area under the receiver operating characteristic; XGBoost, extreme gradient boosting; CatBoost, categorical boosting; SVM, support vector machine

Discussion

The coexistence of diabetes and obesity is prevalent, and sarcopenia poses a serious complication in this population. Prior research indicated that LMM may be a unifying mechanism linking diabetes and obesity [8]. In this study, the prevalence of LMM in patients with obesity and diabetes reached 21.3%, emphasizing its substantial burden. Notably, while patients with LMM are generally thought to have a lower BMI, our study revealed that the LMM group had a significantly higher BMI than the NMM group. This finding aligns with a Chinese study

showing that individuals who were overweight or obese with LMM had higher BMI than those without LMM [30]. The muscle mass deficit is likely masked by fat accumulation, underscoring the limitations of BMI in assessing LMM risk in populations with obesity.

This study is the first to explore the association between novel obesity indices and LMM in patients with obesity and we also developed an online predictive calculator for clinical application. These findings demonstrate a robust association between BRI and LMM, supporting BRI's role as a reliable measure of visceral fat accumulation [20].

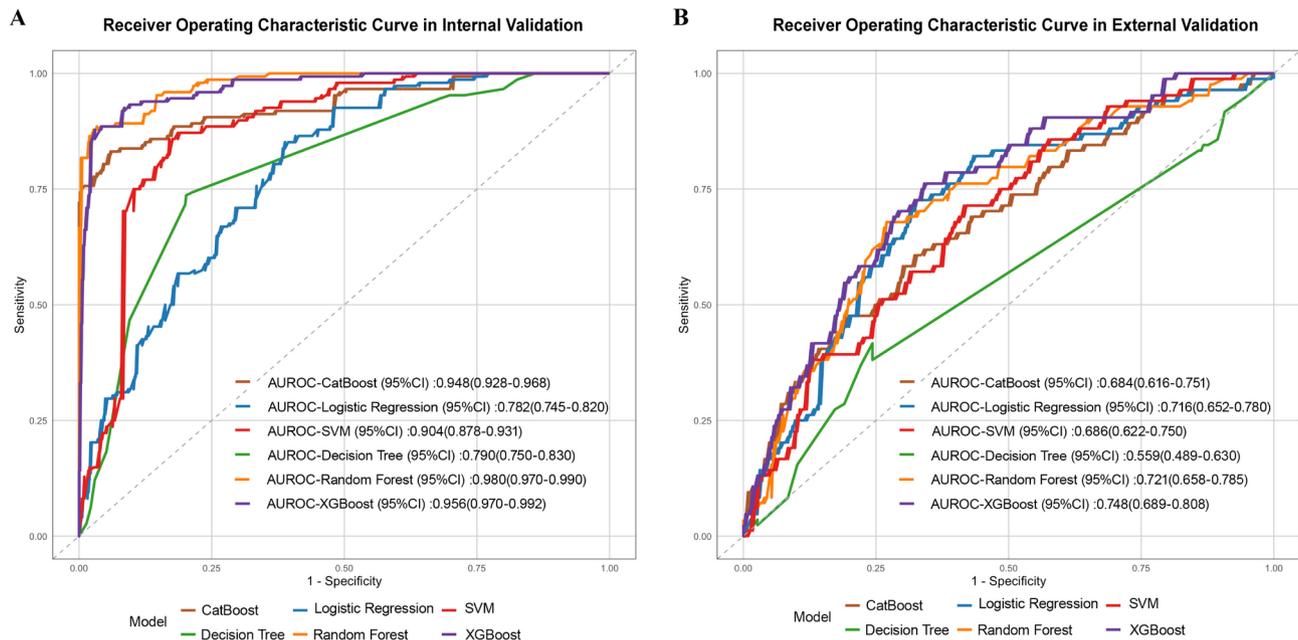


Fig. 3 Receiver operating characteristic curve in internal and external validation sets. **(A)** Receiver operating characteristic curve in internal validation set. **(B)** Receiver operating characteristic curve in external validation set. Abbreviations: CatBoost, categorical boosting; SVM, support vector machine; XGBoost, extreme gradient boosting

Table 3 Key performance indicators of internal and external validation sets

Models	Accuracy	Precision	Recall	F1 score
Internal validation				
Random Forest	0.96	0.93	0.89	0.91
CatBoost	0.92	0.88	0.78	0.82
SVM	0.89	0.73	0.83	0.78
XGBoost	0.87	0.65	0.91	0.76
Decision Tree	0.77	0.50	0.73	0.59
Logistic Regression	0.64	0.38	0.88	0.53
External validation				
Random Forest	0.72	0.48	0.68	0.56
CatBoost	0.66	0.41	0.61	0.49
SVM	0.62	0.39	0.71	0.50
XGBoost	0.69	0.45	0.76	0.57
Decision Tree	0.67	0.39	0.42	0.40
Logistic Regression	0.69	0.46	0.73	0.56

Abbreviations: CatBoost, categorical boosting; SVM, support vector machine; XGBoost, extreme gradient boosting

Visceral fat accumulation can trigger chronic inflammation and insulin resistance, potentially accelerating muscle loss [31, 32]. Although RFM, which estimates total body fat percentage [17], was also positively associated with LMM, the correlation was weaker, suggesting that visceral fat may exert a stronger influence. By contrast, the C-index, an indicator of fat distribution [21], which increased with abdominal fat, did not show a significant association with LMM, indicating that visceral fat accumulation may be more important than overall fat

distribution in the development of LMM among patients with obesity and diabetes.

In the external validation, random forest and extreme gradient boosting (XGBoost) models outperformed other machine learning algorithms, with AUROCs of 0.721 (95%CI: 0.658–0.785) and 0.748 (95%CI: 0.689–0.808), respectively. The random forest model demonstrated outstanding recall and precision, reflecting strong discrimination, accuracy, and generalization. Its F1 score was 0.56, possibly due to differences in the distribution and sample size of training and testing datasets. Nonetheless, decision curve analysis showed a considerable net benefit for diagnosing LMM across several threshold probabilities, and the Brier score of 0.071 highlighted the model’s strong calibration. These results suggest that the random forest model can be a valuable diagnostic tool for LMM in patients with obesity and diabetes and may facilitate clinical decision-making.

The SHAP values further elucidate each variable’s role in the random forest model. BRI identified as the strongest predictor, emphasizing its strong association with muscle loss. Creatinine levels are traditionally considered to be positively correlated with muscle mass, with higher levels typically indicating greater muscle mass [33]. Similarly, our analysis revealed that higher creatinine levels were negatively associated with the prevalence of LMM, suggesting a protective role in model. Age was also a key contributor, suggesting that advanced age is closely associated with an increased risk of LMM. Interestingly, HbA1c displayed a negative association with

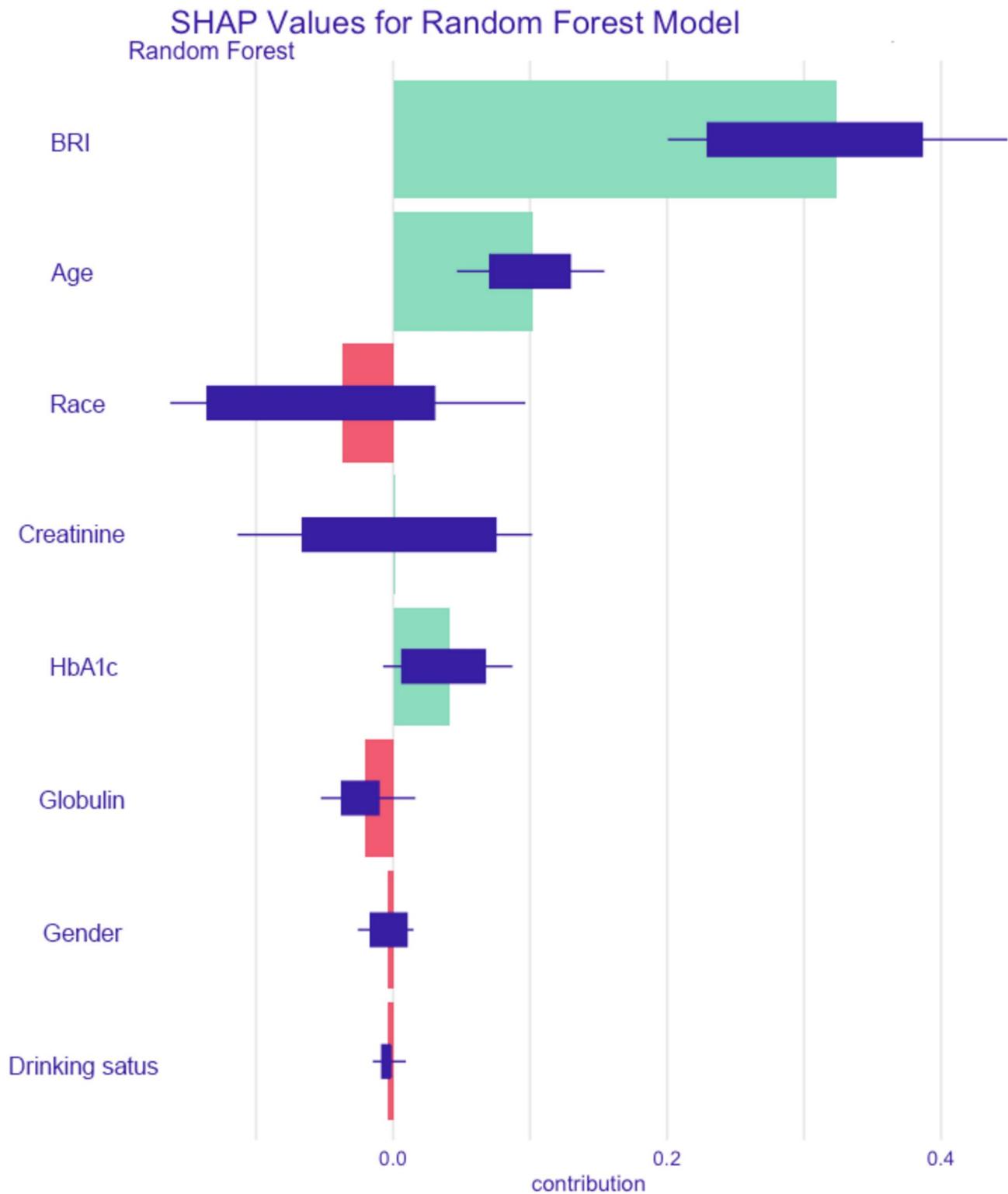


Fig. 4 Feature importance in terms of random forest model

Table 4 Risk stratification of patients in the external validation cohort based on the random forest model

Risk group	Patients(n = 310)	Predicted risk	Observed risk	Pvalue
Low risk (≤ 0.305)	171	14.38%	14.04%	< 0.001
High risk (> 0.305)	139	54.89%	43.17%	

LMM, contrasting with prior evidence that links inadequate glycemic control to muscle loss. For example, a cross-sectional study of older adults in Brazil identified an association between increased blood glucose levels and reduced skeletal muscle mass [34]. Similarly, another study suggested that higher HbA1c levels are associated with an increased risk of LMM, suggesting that poor glycemic control may accelerate muscle loss [35]. Although our analysis considered the use of metformin and insulin, we were unable to fully capture the presence or dosage of additional glucose-lowering agents. It is also noteworthy that individuals with LMM had a significantly higher BMI than those with normal muscle mass (37.80 vs. 33.70 kg/m², $P < 0.001$), which could lead to more aggressive pharmacotherapy in clinical settings. Consistent with this possibility, fasting glucose levels were lower in the LMM group (7.12 vs. 8.99 mmol/L), suggesting intensified glycemic management—even though this difference did not reach statistical significance ($P < 0.100$). Future longitudinal studies with comprehensive medication records and treatment durations will be crucial for clarifying this complex relationship. Additional factors such as globulin levels, race, and sex, also shaped the model. For instance, higher globulin levels may indicate a better nutritional status, contributing to the maintenance of muscle mass. Research has suggested that lower globulin levels are linked to malnutrition and chronic inflammation [36, 37]. Racial disparities may also affect muscle mass and metabolic health; African Americans, for instance, tend to have higher skeletal muscle mass than other ethnic groups [38]. Notably, the variables used in this model are relatively straightforward to assess in routine clinical practice, reinforcing its practical utility.

Prior research indicated that machine learning is effective for predicting LMM or sarcopenia. Kim et al., for instance, used ophthalmic examinations and demographic factors to predict sarcopenia with an AUROC of approximately 0.74 using XGBoost and logistic regression [39]. To our knowledge, this is the first study to construct a predictive model for LMM specifically in patients with obesity and diabetes. The random forest model demonstrated superior discrimination and calibration, with significant differences in the prevalence of LMM between high-risk and low-risk groups (43.17% vs. 14.04%, $P < 0.001$). Notably, the predicted risk values in the high-risk and low-risk group (54.89% vs. 14.38%) were consistent with the observed trend, confirming the model's utility in risk stratification. The high-risk group exhibited a 3.07-fold higher observed risk of LMM compared to the low-risk group, highlighting its potential for guiding clinical decision-making. However, despite comparable baseline characteristics in the training and validation sets ($P > 0.050$, Table S3), the high-risk group showed an 11.7% overestimation of LMM risk (54.89%

predicted vs. 43.17% observed). This discrepancy may be attributable to the relatively small size of the external validation cohort, which can amplify minor calibration errors, as well as uncounted factors (e.g., disease duration, medication dose) that may disproportionately affect the highest-risk subgroup. Crucially, the model's overall discrimination remained robust, indicating that while the absolute predicted probabilities in the high-risk group may be inflated, its capacity to distinguish high-risk and low-risk patients was preserved.

Integrating practical muscle function assessments, such as handgrip strength or gait speed, with the LMM risk prediction model could streamline sarcopenia diagnosis. The cost-effective evaluations reduce reliance on advanced imaging like DXA or MRI, thus enhancing accessibility in low-resource settings without compromising diagnostic accuracy. Despite the promising findings of this study, it had some limitations. First, the external validation cohort, derived via temporal segmentation of the same survey, is smaller than the development cohort, which may limit statistical power. Although this approach provides initial evidence of generalizability, larger and truly independent cohorts will be needed to robustly confirm external validity. Second, to preserve simplicity and enhance generalizability, our model focused on a subset of key predictors, excluding factors such as physical activity and dietary intake. We acknowledge their importance and plan to incorporate them in future large-scale external validation studies to further refine and validate the model. Finally, given the cross-sectional design, establishing causal relationships remains challenging. Future longitudinal studies will be necessary to assess whether the identified risk factors truly predict LMM progression over time.

Conclusions

Our research has significant clinical implications, especially considering the difficulties in diagnosing LMM in patients with obesity and diabetes. The random forest model offers a practical and reliable method for predicting LMM, providing clinicians with an evidence-based tool to guide early risk stratification and management strategies. Moreover, the web-based calculator presents a novel medium for clinical application and generalization of the predictive model.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12944-025-02577-8>.

Supplementary Material 1

Author contributions

The idea for the study was developed by Hui Sheng, who also led the research. Jiaying Ge handled the analyses, result interpretation, and manuscript

preparation. Data collection was performed by Siqi Sun, Jiangping Zeng, Yujie Jing, and Huihui Ma, with oversight and analytical support from Ran Cui, Chunhua Qian, and Shen Qu. All authors reviewed the manuscript.

Funding

This work was supported by the Research Physician Project of Shanghai Tenth People's Hospital (2023YJXSA014) and the National Natural Science Foundation of China (No. 82170894).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 4 March 2025 / Accepted: 18 April 2025

Published online: 29 April 2025

References

1. Koye DN, Magliano DJ, Nelson RG, Pavkov ME. The global epidemiology of diabetes and kidney disease. *Adv Chronic Kidney Dis.* 2018;25:121–32. <https://doi.org/10.1053/j.ackd.2017.10.011>
2. Son JW, Lee SS, Kim SR, Yoo SJ, Cha BY, Son HY, Cho NH. Low muscle mass and risk of type 2 diabetes in Middle-Aged and older adults: findings from the KoGES. *Diabetologia.* 2017;60:865–72. <https://doi.org/10.1007/s00125-016-4196-9>
3. Trierweiler H, Kisielewicz G, Hoffmann Jonasson T, Rasmussen Petterle R, Aguiar Moreira C, Zeghibi Cochenski Borba, V. Sarcopenia: A chronic complication of type 2 diabetes mellitus. *Diabetol Metab Syndr.* 2018;10:25. <https://doi.org/10.1186/s13098-018-0326-5>
4. Filippin LI, Teixeira VN, de Silva O, Miraglia MPM, da Silva F, Sarcopenia FS. A predictor of mortality and the need for early diagnosis and intervention. *Aging Clin Exp Res.* 2015;27:249–54. <https://doi.org/10.1007/s40520-014-0281-4>
5. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. *Lancet.* 2019;393:2636–46. [https://doi.org/10.1016/S0140-6736\(19\)31138-9](https://doi.org/10.1016/S0140-6736(19)31138-9)
6. Licciani AP, Malmstrom TK. Frailty and sarcopenia as predictors of adverse health outcomes in persons with diabetes mellitus. *J Am Med Dir Assoc.* 2016;17:846–51. <https://doi.org/10.1016/j.jamda.2016.07.007>
7. Izzo A, Massimino E, Riccardi G, Della Pepa G. A narrative review on sarcopenia in type 2 diabetes mellitus: prevalence and associated factors. *Nutrients.* 2021;13. <https://doi.org/10.3390/nu13010183>
8. Wolfe RR. The underappreciated role of muscle in health and disease. *Am J Clin Nutr.* 2006;84:475–82. <https://doi.org/10.1093/ajcn/84.3.475>
9. Fabbri E, Chia CW, Spencer RG, Fishbein KW, Reiter DA, Cameron D, Zane AC, Moore ZA, Gonzalez-Freire M, Zoli M, et al. Insulin resistance is associated with reduced mitochondrial oxidative capacity measured by 31P-Magnetic resonance spectroscopy in participants without diabetes from the Baltimore longitudinal study of aging. *Diabetes.* 2017;66:170–6. <https://doi.org/10.2337/db16-0754>
10. Yuan S, Larsson SC. An atlas on risk factors for type 2 diabetes: A Wide-Angled Mendelian randomisation study. *Diabetologia.* 2020;63:2359–71. <https://doi.org/10.1007/s00125-020-05253-x>
11. Visser M, Goodpaster BH, Kritchevsky SB, Newman AB, Nevitt M, Rubin SM, Simonsick EM, Harris TB. Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in Well-Functioning older persons. *J Gerontol Biol Sci Med Sci.* 2005;60:324–33. <https://doi.org/10.1093/gerona/60.3.324>
12. Wang H, Hai S, Liu YX, Cao L, Liu Y, Liu P, Yang Y, Dong BR. Associations between sarcopenic obesity and cognitive impairment in elderly Chinese Community-Dwelling individuals. *J Nutr Health Aging.* 2019;23:14–20. <https://doi.org/10.1007/s12603-018-1088-3>
13. Kong HH, Won CW, Kim W. Effect of sarcopenic obesity on deterioration of physical function in the elderly. *Arch Gerontol Geriatr.* 2020;89:104065. <https://doi.org/10.1016/j.archger.2020.104065>
14. Saito H, Matsue Y, Kamiya K, Kagiyaama N, Maeda D, Endo Y, Ueno H, Yoshioka K, Mizukami A, Saito K, et al. Sarcopenic obesity is associated with impaired physical function and mortality in older patients with heart failure: insight from FRAGILE-HF. *BMC Geriatr.* 2022;22:556. <https://doi.org/10.1186/s12877-022-03168-3>
15. Ge J, Zeng J, Li N, Ma H, Zhao Z, Sun S, Jing Y, Qian C, Fei Z, Qu S, et al. Soluble Interleukin 2 receptor is risk for sarcopenia in men with high fracture risk. *J Orthop Translat.* 2022;38:213–9. <https://doi.org/10.1016/j.jtot.2022.10.017>
16. Ge J, Zeng J, Ma H, Sun S, Zhao Z, Jing Y, Qian C, Fei Z, Cui R, Qu S, et al. A new index based on serum creatinine and Cystatin C can predict the risks of sarcopenia, falls and fractures in old patients with low bone mineral density. *Nutrients.* 2022;14:5020. <https://doi.org/10.3390/nu14235020>
17. Suthahar N, Wang K, Zwartkruis VW, Bakker SJL, Inzucchi SE, Meems LMG, Eijgenraam TR, Ahmadizar F, Sijbrands EG, Gansevoort RT, et al. Associations of relative fat mass, a new index of adiposity, with Type-2 diabetes in the general population. *Eur J Intern Med.* 2023;109:73–8. <https://doi.org/10.1016/j.ejim.2022.12.024>
18. Jablonowska-Lietz B, Wrzosek M, Włodarczyk M, Nowicka G. New indexes of body fat distribution, visceral adiposity index, body adiposity index, Waist-to-Height ratio, and metabolic disturbances in the obese. *Kardiol Pol.* 2017;75:1185–91. <https://doi.org/10.5603/KPa.2017.0149>
19. Guo X, Ding Q, Liang M. Evaluation of eight anthropometric indices for identification of metabolic syndrome in adults with diabetes. *Diabetes Metab Syndr Obes.* 2021;14:1431–43. <https://doi.org/10.2147/DMSO.S294244>
20. Thomas DM, Bredlau C, Bony-Westphal A, Mueller M, Shen W, Gallagher D, Maeda Y, McDougall A, Peterson CM, Ravussin E, et al. Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. *Obesity.* 2013;21:2264–71. <https://doi.org/10.1002/oby.20408>
21. Ruperto M, Barril G, Sánchez-Muniz FJ. Usefulness of the conicity index together with the conjoint use of adipocytokines and Nutritional-Inflammatory markers in Hemodialysis patients. *J Physiol Biochem.* 2017;73:67–75. <https://doi.org/10.1007/s13105-016-0525-1>
22. Wang D, Chen Z, Wu Y, Ren J, Shen D, Hu G, Mao C. Association between two novel anthropometric measures and type 2 diabetes in a Chinese population. *Diabetes Obes Metabolism.* 2024;26:3238–47. <https://doi.org/10.1111/dom.15651>
23. Zhao Q, Zhang K, Li Y, Zhen Q, Shi J, Yu Y, Tao Y, Cheng Y, Liu Y. Capacity of a body shape index and body roundness index to identify diabetes mellitus in Han Chinese people in Northeast China: A Cross-Sectional study. *Diabet Med.* 2018;35:1580–7. <https://doi.org/10.1111/dme.13787>
24. The Relationship of Body Composition. Indices with the significance, extension and severity of coronary artery disease. *Nutr Metabolism Cardiovasc Dis.* 2020;30:2279–85. <https://doi.org/10.1016/j.numecd.2020.07.014>
25. American Diabetes Association Professional Practice Committee 2. Classification and diagnosis of diabetes: standards of medical care in Diabetes-2022. *Diabetes Care.* 2022;45:517–38. <https://doi.org/10.2337/dc22-5002>
26. Batsis JA, Mackenzie TA, Bartels SJ, Sahakyan KR, Somers VK, Lopez-Jimenez F. Diagnostic accuracy of body mass index to identify obesity in older adults: NHANES 1999–2004. *Int J Obes (Lond).* 2016;40:761–7. <https://doi.org/10.1038/ijo.2015.243>
27. Studenski SA, Peters KW, Alley DE, Cawthon PM, McLean RR, Harris TB, Ferrucci L, Guralnik JM, Fragala MS, Kenny AM et al. The FNIH Sarcopenia Project: Rationale, Study Description, Conference Recommendations, and Final Estimates. *J Gerontol A Biol Sci Med Sci.* 2014;69:547–558. <https://doi.org/10.1093/gerona/glu010>
28. Shen L, Xu X, Yue S, Yin SA. Predictive model for depression in Chinese Middle-Aged and elderly people with physical disabilities. *BMC Psychiatry.* 2024;24:305. <https://doi.org/10.1186/s12888-024-05766-4>
29. Kalyani RR, Corriere M, Ferrucci L. Age-Related and Disease-Related muscle loss: the effect of diabetes, obesity, and other diseases. *Lancet Diabetes Endocrinol.* 2014;2:819–29. [https://doi.org/10.1016/S2213-8587\(14\)70034-8](https://doi.org/10.1016/S2213-8587(14)70034-8)
30. Associations between Sarcopenic Obesity and Risk of Cardiovascular Disease. A Population-Based cohort study among Middle-Aged and older adults using the CHARLS. *Clin Nutr.* 2024;43:796–802. <https://doi.org/10.1016/j.clnu.2024.02.002>
31. Barbosa P, Pinho A, Lázaro A, Rosendo-Silva D, Paula D, Campos J, Tralhão JG, Pereira MJ, Paiva A, Laranjeira P, et al. CD8+ Treg cells play a role in the Obesity-Associated insulin resistance. *Life Sci.* 2024;336:122306. <https://doi.org/10.1016/j.lfs.2023.122306>
32. Elkins C, Li C. Deciphering visceral adipose tissue regulatory T cells: key contributors to metabolic health. *Immunol Rev.* 2024;324:52–67. <https://doi.org/10.1111/immr.13336>

33. Groothof D, Shehab NB, Eler NS, Post A, Kremer D, Polinder-Bos HA, Gansevoort RT, Groen H, Pol RA, Gans RO, Creatinine, Cystatin C, et al. Muscle mass, and mortality: findings from a primary and replication Population-based cohort. *J Cachexia Sarcopenia Muscle*. 2024;15:1528. <https://doi.org/10.1002/jcsm.13511>
34. de Campos GC, Lourenço RA, Lopes CS. Prevalence of sarcopenic obesity and its association with functionality, lifestyle, biomarkers and morbidities in older adults: the FIBRA-RJ study of frailty in older Brazilian adults. *Clin (Sao Paulo)*. 2020;75:e1814. <https://doi.org/10.6061/clinics/2020/e1814>
35. Pereira S, Marliss EB, Morais JA, Chevalier S, Gougeon R. Insulin resistance of protein metabolism in type 2 diabetes. *Diabetes*. 2008;57:56–63. <https://doi.org/10.2337/db07-0887>
36. Duvall LE, Shipman AR, Shipman KE. Investigative algorithms for disorders affecting plasma proteins with a focus on albumin and the calculated Globulin fraction: A narrative review. *J Lab Precision Med*. 2023;8. <https://doi.org/10.21037/jlpm-23-15>
37. Wang X, Zhang J, Xu X, Pan S, Cheng L, Dang K, Qi X, Li Y. Associations of daily eating frequency and nighttime fasting duration with biological aging in National health and nutrition examination survey (NHANES) 2003–2010 and 2015–2018. *Int J Behav Nutr Phys Act*. 2024;21:104. <https://doi.org/10.1186/s12966-024-01654-y>
38. Lopes LCC, Vaz-Gonçalves L, Schincaglia RM, Gonzalez MC, Prado CM, Oliveira EP. Sex and Population-Specific cutoff values of muscle quality index: results from NHANES 2011–2014. *Clin Nutr*. 2022;41:1328–34. <https://doi.org/10.1016/j.clnu.2022.04.026>. MotaJ.F.
39. Kim BR, Yoo TK, Kim HK, Ryu IH, Kim JK, Lee IS, Kim JS, Shin D-H, Kim Y-S, Kim BT. Oculomics for sarcopenia prediction: A machine learning approach toward predictive, preventive, and personalized medicine. *EPMA J*. 2022;13:367–82. <https://doi.org/10.1007/s13167-022-00292-3>

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.