

Multiparametric Rectus Femoris Ultrasound With and Without Serum Biomarkers for Identifying Sarcopenia in Type 2 Diabetes Mellitus: A Cross-Sectional Model-Development Study

Yong He

Wuhu

Mengying Wu

Wuhu

Zhao Sheng

Wuhu

Daojing Xu

Wuhu

Shaoyang Guan

Wuhu

Jiawei Fan

Wuhu

Jinfeng Ma

13909631027@163.com

The Second Affiliated Hospital of Wannan Medical University

Article

Keywords: sarcopenia, type 2 diabetes mellitus, shear-wave elastography, triglyceride-glucose index, systemic immune-inflammation index, model development, internal validation

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10986988/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations:

No competing interests reported.

Tables are available in the Supplementary Files section.

Multiparametric Rectus Femoris Ultrasound With and Without Serum Biomarkers for Identifying Sarcopenia in Type 2 Diabetes Mellitus: A Cross-Sectional Model-Development Study

Yong He^{1#}, Mengying Wu^{2#}, Zhao Sheng³, Daojing Xu¹, Shaoyang Guan³, Jiawei Fan^{4*}, Jinfeng Ma^{5*}

¹Department of Ultrasound Medicine, The Second People's Hospital, Wuhu, Wuhu, Anhui, China

²Department of Pharmacy, The Second People's Hospital, Wuhu, Wuhu, Anhui, China

³Department of Endocrinology, The Second People's Hospital, Wuhu, Wuhu, Anhui, China

⁴Department of Geriatrics, The Second People's Hospital, Wuhu, Wuhu, Anhui, China

⁵Department of Gynecology, The Second Affiliated Hospital of Wannan Medical University, Wuhu, Anhui, China

These authors contributed equally to this work: Yong He and Mengying Wu.

*Correspondence: Jinfeng Ma, 13909631027@163.com; Jiawei Fan, wheyfjw@163.com

Abstract

Among 209 hospitalized patients with type 2 diabetes mellitus (T2DM), including 65 with sarcopenia defined by the 2025 Asian Working Group for Sarcopenia criteria, we evaluated an age-adjusted multiparametric rectus femoris ultrasound model and compared it with serum and combined models. Ultrasound variables were rectus femoris thickness, echogenicity, and shear-wave elastography; serum variables were the triglyceride-glucose (TyG) and systemic immune-inflammation (SII) indices. The age-adjusted ultrasound model had an apparent AUC of 0.986 (95% CI 0.976-0.997), an optimism-corrected calibration slope of 0.825, and a cross-validated AUC of 0.983 (SD 0.005). The age-only AUC was 0.796; adding the ultrasound variables increased the AUC by 0.190 (DeLong $P = 5.0 \times 10^{-11}$). Among participants aged 65 years or older ($n = 120$), the age-only AUC was 0.603. Adding TyG and SII increased the apparent AUC from 0.986 to 0.995 (DeLong $P = 0.045$), but categorical reclassification was not significant ($P = 0.106$), decision-curve gains were small, and the combined-model optimism-corrected calibration slope was 0.484. After removal of 27 observations flagged by influence diagnostics, conventional estimation of the combined model did not yield finite estimates for several coefficients; in the full cohort, Firth penalized estimation yielded a TyG_Z odds ratio of 10.86 (95% CI 2.59-88.33). The findings support further evaluation of multiparametric rectus femoris ultrasound as a candidate case-finding approach in this setting. The incremental role of TyG and SII remains uncertain, and external validation is required before clinical use.

Keywords: sarcopenia; type 2 diabetes mellitus; shear-wave elastography; triglyceride-glucose index; systemic immune-inflammation index; model development; internal validation

Introduction

Sarcopenia, characterized by a progressive decline in skeletal muscle mass and function, was recognized as an independent disease entity in the International Classification of Diseases, Tenth Revision, in 2016 ¹. Its clinical consequences include increased risks of falls, fragility fractures, prolonged hospitalization, and premature mortality ². Sarcopenia is particularly common among patients with type 2 diabetes mellitus (T2DM), in whom metabolic and inflammatory abnormalities converge to impair muscle health. Recent meta-analyses in Asian adults with T2DM reported pooled sarcopenia prevalence estimates of 17% and 23%, reflecting differences in included studies and diagnostic definitions ^{3,4}. This elevated burden highlights the need for accurate case identification in patients with diabetes.

The 2025 AWGS consensus update introduced important refinements to the diagnostic framework ⁵. It emphasizes early promotion of muscle health and provides diagnostic thresholds for adults aged 50-64 years as well as those aged 65 years or older. Sarcopenia requires both low muscle mass and low muscle strength, whereas physical performance measures, including gait speed and the five-time chair stand test, are classified as functional outcomes rather than mandatory diagnostic components. Dual-energy X-ray absorptiometry (DXA) and bioelectrical impedance analysis (BIA) remain standard methods for quantifying muscle mass, but both have limitations. DXA requires fixed equipment, whereas BIA estimates are sensitive to hydration status and depend on population- and device-specific prediction equations ^{1,6}. Moreover, neither method distinguishes contractile tissue from intramuscular fat infiltration (myosteatosis) or assesses muscle mechanical properties. These limitations have increased interest in neuromuscular ultrasound as a point-of-care assessment tool ⁷.

High-resolution ultrasound provides noninvasive, radiation-free assessment of skeletal muscle ⁷. Morphologic measures such as muscle thickness can be complemented by the echogenicity index (EI), which has been associated with fibro-adipose infiltration ⁸. Shear-wave elastography (SWE) quantifies tissue stiffness and may provide mechanical information related to muscle structure and function ⁹. Previous studies have reported associations between rectus femoris SWE and sarcopenia-related phenotypes beyond conventional grayscale B-mode measures ^{12,13}. Combining thickness, echogenicity, and stiffness may therefore capture complementary structural and mechanical characteristics of muscle ¹⁴.

Circulating markers of insulin resistance and inflammation have also been associated with sarcopenia. The triglyceride-glucose (TyG) index, derived from fasting glucose and triglyceride concentrations, is a practical surrogate of insulin resistance ¹⁵. A meta-analysis of 20 cross-sectional studies involving 521,778 participants reported higher odds of sarcopenia with higher TyG values (OR = 1.63, 95% CI: 1.29-2.07) ¹⁶. Proposed mechanisms include impaired insulin signaling, reduced muscle

protein synthesis, and anabolic resistance 17. The systemic immune-inflammation index (SII), calculated from platelet, neutrophil, and lymphocyte counts, summarizes systemic inflammatory status 18. A 2024 meta-analysis found higher SII levels in adults with sarcopenia, although estimates comparing high versus low SII were heterogeneous 19. Both indices are inexpensive and routinely available, making their relation to sarcopenia in T2DM clinically relevant to examine.

Whether TyG and SII add useful information to a multiparametric ultrasound model is less clear. Ultrasound and circulating biomarkers may reflect partly overlapping aspects of the metabolic, inflammatory, structural, and mechanical changes that accompany muscle loss. Few studies have quantified the added value of TyG or SII beyond advanced ultrasound in T2DM, particularly using reclassification and decision-curve measures rather than AUC alone.

We therefore evaluated two questions in hospitalized patients with T2DM: how well an age-adjusted multiparametric rectus femoris ultrasound model discriminates AWGS 2025-defined sarcopenia, with muscle mass measured by DXA; and whether TyG and SII add information beyond the ultrasound model. Because the ultrasound model showed high discrimination in this cohort, the biomarker comparison was treated as exploratory.

Methods

Study Design and Participants

This cross-sectional study included hospitalized patients with T2DM at the Second People's Hospital of Wuhu, China, between July 2024 and July 2026. Ultrasound data were available for 297 patients during the study period. The analytical cohort comprised 209 patients who completed rectus femoris ultrasound, DXA, functional assessment, and the prespecified analyses. Enrolment was not prospectively logged in a form that permits verification of consecutive entry; possible selection and spectrum effects are addressed in the Limitations.

The inclusion criteria were as follows: (1) T2DM diagnosed according to the Chinese Guidelines for the Prevention and Treatment of Diabetes Mellitus (2024 Edition) 20; (2) age ≥ 50 years; (3) sufficient physical and cognitive ability to complete ultrasound, dual-energy X-ray absorptiometry (DXA), and functional performance assessments; and (4) complete data for the prespecified primary model variables and reference-standard classification.

The exclusion criteria were as follows: (1) severe acute infection, cardiopulmonary insufficiency, or severe hepatic or renal dysfunction; (2) major musculoskeletal trauma, orthopedic surgery, or active malignancy within the preceding 6 months; (3) severe psychiatric or neurologic disorders that precluded

independent activities of daily living; (4) chronic use of systemic glucocorticoids or other medications known to substantially affect skeletal muscle metabolism; and (5) a previous diagnosis of sarcopenia or current participation in a structured rehabilitation program targeting sarcopenia.

The pathway from data collection to the analytical cohort is shown in Figure 1. Ultrasound measurements were obtained in 297 patients. Of these, 88 were excluded: 42 did not complete DXA, 28 did not complete the handgrip, 6-m gait speed, or five-time chair stand assessments, and 18 had incomplete prespecified laboratory assessments, for example the biochemistry panel or the insulin and C-peptide measurements. The remaining 209 patients formed the analytical cohort. These counts were reconstructed retrospectively from retained electronic medical record data and do not substitute for a prospective screening log.

The study protocol was approved by the Institutional Ethics Committee of the Second People's Hospital of Wuhu (approval no. 2024-KY-231). All participants provided written informed consent before enrollment. The study was not prospectively registered.

Reporting guideline: This cross-sectional model-development study is reported with reference to STARD 2015 and the STROBE statement for cross-sectional studies. The completed STARD 2015 checklist is provided as a supplementary file.

The age-adjusted multiparametric ultrasound model (Model 1: age, STD-RFthick, SWE_{mean}, and EI) was prespecified as the primary model. A serum-only model (Model 2: age, TyG_Z, and SII_Z) and a combined model (Model 3: all six variables) were evaluated as comparisons. Because conventional estimation of Model 3 proved unstable (see Statistical Analysis), Model 1 serves as the primary model throughout, and Model 3 is reported as an exploratory analysis. Sarcopenia classified according to the 2025 AWGS criteria, based on DXA-derived appendicular muscle mass and handgrip strength, served as the reference standard.

Sarcopenia Assessment

Sarcopenia was classified according to the 2025 AWGS consensus update 5, requiring low appendicular muscle mass and low muscle strength. Outcome classification was verified against the age- and sex-specific thresholds described below for all 209 participants.

Muscle Mass: Appendicular skeletal muscle mass was measured using a Hologic Horizon W whole-body DXA system. Before scanning, participants were instructed to fast overnight, empty their bladder and bowels, remove all metal objects, and wear standardized light clothing. A certified technician performed all scans according to the manufacturer's protocol. The appendicular skeletal muscle mass index (ASMI) was calculated as appendicular muscle mass (kg) divided by height squared (m²). Height-

adjusted ASMI was used because the prespecified reference-standard definition applied the explicit age- and sex-specific DXA cutoffs used in the 2025 AWGS framework. Alternative muscle-mass adjustments were not evaluated and are considered in the Limitations. Low muscle mass was defined as $< 7.0 \text{ kg/m}^2$ for men and $< 5.4 \text{ kg/m}^2$ for women aged ≥ 65 years, and $< 7.2 \text{ kg/m}^2$ for men and $< 5.5 \text{ kg/m}^2$ for women aged 50-64 years.

Muscle Strength: Maximum isometric handgrip strength was measured using a Jamar hydraulic hand dynamometer (Sammons Preston, Bolingbrook, IL, USA). Participants were seated with the shoulder adducted and neutrally rotated, the elbow flexed to 90 degrees, and the forearm in a neutral position. Three maximal-effort measurements were obtained from each hand in alternating order, with 60 seconds of rest between attempts. The highest value from the six measurements was recorded. Low muscle strength was defined as $< 28 \text{ kg}$ for men and $< 18 \text{ kg}$ for women aged ≥ 65 years, and $< 34 \text{ kg}$ for men and $< 20 \text{ kg}$ for women aged 50-64 years.

Physical Performance: Physical performance was assessed using the 6-m usual gait speed test and the five-time chair stand test.

Rectus Femoris Ultrasound Assessment

All ultrasound examinations were performed by a physician with more than five years of experience in musculoskeletal ultrasonography. The ultrasound operator was blinded to the participants' clinical data, DXA results, and sarcopenia status. Blinding of the personnel performing DXA and the functional assessments to the ultrasound results was not separately recorded in the retained study documentation. Imaging was performed using a Siemens ACUSON Sequoia system equipped with a 10L4 linear-array transducer (3-10 MHz) and two-dimensional color shear-wave elastography (SWE).

B-mode Morphologic Measurements: Participants rested in the supine position for 15 minutes with their lower limbs fully relaxed. The transducer was placed perpendicular to the skin and to the long axis of the limb, producing a transverse rectus femoris section at a point three-fifths of the distance from the anterior superior iliac spine to the superior border of the patella on the dominant side. This transverse plane aligns the shear-wave propagation axis obliquely to the longitudinal course of the rectus femoris fibers; the implications of this orientation are discussed in the Limitations. After an optimal transverse image had been stabilized, the region of interest (ROI) was manually traced along the outer epimysial border of the rectus femoris to measure muscle thickness (RFthick), circumference (RFc), and cross-sectional area (RFcsa). To account for body size, these measurements were standardized to height and expressed as STD-RFthick (mm/m), STD-RFc (mm/m), and STD-RFcsa (cm²/m).

B-mode and SWE acquisition parameters were fixed for all participants and recorded at the start of each session: B-mode center frequency 8 MHz within the 3-10 MHz range of the 10L4 transducer; depth 4.0 cm; a single focus positioned at the level of the rectus femoris; gain adjusted at protocol standardization so that the subcutaneous fat layer displayed mid-range echogenicity, then held constant; dynamic range 60 dB; time-gain compensation set to a linear ramp and not adjusted thereafter; and SWE display scale fixed at 0-100 kPa.

Shear-Wave Elastography (SWE): Two-dimensional color elastograms were acquired at the same anatomical site while the muscle was fully relaxed. Five circular ROIs, each 5 mm in diameter, were placed in the four quadrants and center of the rectus femoris cross-section. Shear-wave velocities were automatically converted to elastic modulus values (kPa) by the system software. Each ROI was sampled three times, yielding 15 measurements per participant; the mean of these measurements was recorded as SWE_{mean}. An acquisition was accepted only when the color fill was judged complete across all five ROIs over five consecutive frames and the system quality indicator was displayed as high; acquisitions were repeated if the color fill was incomplete, if the quality indicator fell below the required level, or if the participant moved. Joint position, the minimal-compression technique, and the rejection and repetition rules were prespecified and recorded.

Echogenicity Index (EI): Raw DICOM images were exported and analyzed using ImageJ software (version 1.53; National Institutes of Health, Bethesda, MD, USA). An ROI was manually delineated along the inner echogenic border of the rectus femoris fascia, excluding fascia, subcutaneous fat, macroscopic blood vessels, and focal calcifications. Mean pixel intensity, ranging from 0 to 255 arbitrary units (AU), was obtained from the grayscale histogram. Higher EI indicates greater echogenicity and has been associated with fibro-adipose replacement. Three measurements were obtained and averaged. Because grayscale intensity depends strongly on machine settings, the EI values reported here are specific to this system and fixed preset and require recalibration before comparison with other equipment or settings.

Reproducibility Analysis

To assess measurement reliability, 120 participants were allocated to two prespecified groups (60 per group). Group 0 assessed intra-observer/test-retest repeatability: the primary operator repeated all measurements after 48 hours while blinded to the initial results. Group 1 assessed inter-observer agreement: a second operator repeated the examination using the same anatomical landmark, participant position, ultrasound settings, ROI definition, and minimal-compression protocol while blinded to the primary operator's results and clinical status. ICC(2,1), a two-way random-effects, absolute-agreement, single-measurement model, was calculated separately for each group, with bootstrap 95% intervals from

2,000 resamples. Repeat measurements were obtained from newly acquired images, so the ICCs reflect both acquisition and measurement reproducibility.

Laboratory Measurements and Biomarker Calculation

Fasting venous blood samples were collected after an overnight fast of at least 8 hours. Fasting plasma glucose (FPG), triglycerides (TG), total cholesterol, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol were measured using standard automated biochemical analyzers. Complete blood counts were obtained using an automated hematology analyzer.

The TyG index was calculated as $\ln [\text{fasting TG (mg/dL)} \times \text{FPG (mg/dL)} / 2]$. When laboratory values were reported in mmol/L, the following conversion factors were applied before calculation: $\text{TG (mg/dL)} = \text{TG (mmol/L)} \times 88.57$ and $\text{FPG (mg/dL)} = \text{FPG (mmol/L)} \times 18.00$.

SII was calculated as $\text{platelet count (109/L)} \times \text{neutrophil count (109/L)} / \text{lymphocyte count (109/L)}$.

To permit direct comparison of effect estimates in the regression models, the TyG index and SII were standardized as Z scores (TyG_Z and SII_Z) using the mean and standard deviation of the study cohort.

Diabetes treatment regimen was not captured as a structured study variable and could not be reconstructed reliably for all participants from narrative discharge records. We therefore did not create a treatment variable; this gap is acknowledged in the Limitations.

Statistical Analysis

The final cohort included 209 patients and 65 sarcopenia events. All primary model variables were complete. However, HbA1c was missing in 7 participants, albumin in 4, HDL-C in 2, and LDL-C in 1; these variables were not used in the primary models. The event-per-variable ratio was approximately 16.3 for the four-predictor primary model and approximately 10.8 for the six-predictor combined model. The study size reflected the number of eligible participants available during the study period, and the limited number of events was considered when restricting the number of model parameters and interpreting model stability.

Data verification. The audited dataset contained 209 patients and 41 variables. All six candidate model variables were complete, so no imputation was used. Variable ranges were checked against physiological plausibility, the TyG and SII indices were recomputed from their constituent laboratory values, and the Z-standardized values were confirmed to reproduce from the raw indices. No values were edited manually, and no observations were excluded on the basis of value plausibility.

The distribution of continuous variables was assessed using the Shapiro-Wilk test. Normally distributed variables are presented as mean +/- standard deviation (SD), and non-normally distributed

variables as median (interquartile range [IQR]). Categorical variables are presented as number (percentage). Between-group comparisons were performed using the independent-samples t test or Mann-Whitney U test for continuous variables and the chi-square test or Fisher's exact test for categorical variables, as appropriate.

Univariable and multivariable conventional maximum-likelihood logistic regression analyses were performed. The six variables used in the prespecified models were selected from the study question and prior literature rather than from baseline significance testing. Model 1 included age, STD-RFthick, SWEmean, and EI; Model 2 included age, TyG_Z, and SII_Z; and Model 3 included all six variables. STD-RFcsa was excluded because of its correlation with thickness and the limited number of events. Sex and BMI were not included in the primary specifications to limit model dimensionality; sensitivity models added sex, BMI, and both variables. Because sparse data and near-separation can inflate maximum-likelihood coefficients, Model 3 was also fitted using Firth penalized-likelihood logistic regression with penalized profile-likelihood 95% confidence intervals. Influence diagnostics were added after the instability of Model 3 became apparent. Discrimination, calibration, reclassification, and decision-curve analyses for the primary ultrasound model remained based on the conventional logistic model.

Regression diagnostics were performed for all multivariable models. Restricted-cubic-spline sensitivity analyses did not support significant nonlinearity for EI (chi-square = 4.47 on 2 df; $P = 0.107$) or SII_Z (chi-square = 2.69 on 2 df; $P = 0.261$); the observed decile patterns are reported descriptively in the Results. Box-Tidwell tests did not detect nonlinearity for age, STD-RFthick, SWEmean, or TyG_Z. VIF values were below 2.0 for the ultrasound and serum models and reached 5.04 in the combined model. Influence diagnostics for the combined model used Cook's distance $> 4/n$ and leverage $> 3p/n$; 14 and 26 observations exceeded these thresholds, respectively, with 27 participants flagged by either criterion. Individual diagnostic values are provided in Supplementary Table S3.

Discrimination was quantified by AUC with DeLong confidence intervals. Sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and positive and negative likelihood ratios were calculated at cohort-derived Youden cutoffs with 95% confidence intervals. Because no external cutoff was prespecified and the cutoffs were selected and evaluated in the same cohort, these classification measures are presented as apparent descriptive estimates rather than internally validated operating points. Internal validation therefore focused on threshold-independent discrimination, calibration slope, and Brier score. Model fit was compared using AIC and likelihood-ratio tests. Reclassification and decision-curve analyses were secondary analyses. No imputation was required for the primary models.

Internal validation used repeated 10-fold cross-validation with 100 repetitions and bootstrap optimism correction with 1,000 resamples. Bootstrap-corrected AUC, calibration slope, and Brier score were reported, and bootstrap optimism-corrected calibration curves were generated for Models 1-3. Calibration-in-the-large was also examined but was not used as evidence of external calibration.

Categorical NRI was evaluated at 10% and 15% thresholds, and decision-curve analysis was examined across threshold probabilities of 5-25%. These thresholds were chosen for exploratory comparison and were not tied to a prespecified clinical management action.

Sensitivity analyses. Sex, BMI, and both were added to the combined model to assess their effect on model performance (Supplementary Table S4).

All analyses were performed using R software (version 4.6.1; R Foundation for Statistical Computing, Vienna, Austria) with the rms, pROC, PredictABEL, dcurves, ResourceSelection, car, and logistf packages. All tests were two-sided, and $P < 0.05$ was considered statistically significant. For the Firth sensitivity analysis, inference was based on penalized profile-likelihood 95% confidence intervals.

Data availability

All relevant data are within the paper. The raw data used and analyzed during the current study are available from the corresponding author upon reasonable request.

Supplementary Information

Supplementary Information: Supplementary Table S1, Supplementary Tables S2-S4, Supplementary Figures S1-S2, and the STARD 2015 checklist accompany this manuscript.

Results

Reproducibility of Ultrasound Measurements

Ultrasound measurements showed good-to-excellent repeatability and inter-observer agreement. Intra-observer ICCs ranged from 0.874 (SWEmean) to 0.922 (STD-RFthick), and inter-observer ICCs from 0.869 (STD-RFcsc) to 0.921 (EI) (Table 2). Because repeat assessments involved newly acquired images, these estimates reflect variability in both image acquisition and measurement.

Study Population and Baseline Characteristics

Ultrasound data were available for 297 hospitalized patients with T2DM during the study period. Of these, 42 did not complete DXA, 28 did not complete the prespecified functional assessments, and 18 had incomplete prespecified laboratory assessments, leaving 209 participants in the analytical cohort (Figure

1). Among these participants, 65 (31.1%) met the 2025 AWGS criteria for sarcopenia. The cohort comprised patients who completed the required assessments during the study period rather than a prospectively verified consecutive series.

Patients with sarcopenia were older than those without sarcopenia (median, 71.0 vs. 62.0 years; $P < 0.001$), and a greater proportion were aged ≥ 65 years (90.8% vs. 42.4%; $P < 0.001$). Median body mass index was lower in the sarcopenia group (23.12 vs. 24.20 kg/m²; $P = 0.017$). Sex distribution, hypertension prevalence, and diabetes duration did not differ significantly between groups. Diabetes treatment regimens were not systematically available and therefore could not be characterized.

Table 1 reports the remaining hematological and biochemical values. On ultrasound, patients with sarcopenia had lower STD-RFthick, STD-RFc_{sa}, and SWEmean and higher EI (all $P < 0.001$), whereas STD-RFc did not differ significantly ($P = 0.233$). TyG and SII were also higher in the sarcopenia group (both $P < 0.001$).

Factors Associated with Sarcopenia

In univariable analyses, age, STD-RFthick, SWEmean, EI, TyG_Z, and SII_Z were each associated with sarcopenia (all $P \leq 0.001$; Table 3). These odds ratios describe cross-sectional associations rather than effects.

In Model 1, age, STD-RFthick, SWEmean, and EI remained associated with sarcopenia; in Model 2, age, TyG_Z, and SII_Z remained associated (Table 3).

The full conventional regression results, including effect estimates and confidence intervals, are presented in Table 3.

In the conventional combined model (Model 3), all six variables were associated with sarcopenia, but the TyG_Z estimate was large and imprecise (OR = 26.490, 95% CI: 3.012-232.982). Influence diagnostics flagged 27 participants. After their removal, conventional maximum-likelihood estimation did not yield finite estimates for TyG_Z, SWEmean, or EI, showing substantial coefficient sensitivity. In the full cohort, Firth penalization produced finite estimates with unchanged directions; TyG_Z was attenuated to OR = 10.86 (95% CI: 2.59-88.33) and SII_Z to OR = 3.40 (95% CI: 1.40-11.65). Full Firth estimates are provided in Supplementary Table S1. Figure 2 and Table 3 show the conventional maximum-likelihood estimates.

Model Discrimination and Comparison

The age-only model had an AUC of 0.796 (95% CI, 0.762-0.830). The age-adjusted ultrasound model (Model 1) had an AUC of 0.986 (95% CI, 0.976-0.997), the serum model (Model 2) had an AUC

of 0.932 (95% CI, 0.900-0.965), and the combined model (Model 3) had an AUC of 0.995 (95% CI, 0.990-1.000). Adding the ultrasound parameters to age increased the AUC by 0.190 (95% CI, 0.178-0.202; DeLong $P = 5.0 \times 10^{-11}$). Removing age from the ultrasound model resulted in an AUC of 0.984. The combined-versus-ultrasound AUC difference was 0.0089 (95% CI, 0.0002-0.0176; DeLong $P = 0.0447$). Diagnostic performance of Models 1-3 and the complete pairwise DeLong comparisons are shown in Table 4; the age-only and no-age sensitivity analyses are summarized with the internal-validation results in Table 5.

At the cohort-derived Youden cutoffs, Model 1 had sensitivity 93.8% (95% CI, 85.0-98.3%) and specificity 93.8% (95% CI, 88.5-97.1%); Model 2 had sensitivity 90.8% (95% CI, 81.0-96.5%) and specificity 84.0% (95% CI, 77.0-89.6%); and Model 3 had sensitivity 96.9% (95% CI, 89.3-99.6%) and specificity 96.5% (95% CI, 92.1-98.9%). Because these cutoffs were derived and evaluated in the same cohort, these estimates are descriptive and are not intended for direct application in external populations. The receiver operating characteristic curves are shown in Figure 3.

Model Fit, Calibration, and Internal Validation

The combined model had a lower AIC (45.73) than the ultrasound model (65.44) and serum model (138.22), and the likelihood-ratio test comparing combined and ultrasound models was significant (chi-square = 23.71, $df = 2$, $P < 0.001$). Its maximum VIF was 5.04 and its optimism-corrected calibration slope was 0.484, indicating considerably greater instability of fitted probabilities than in the primary ultrasound model.

Repeated 10-fold cross-validation with 100 repetitions yielded mean AUCs of 0.794 (SD 0.013) for age alone, 0.983 (SD 0.005) for the age-adjusted ultrasound model, 0.982 (SD 0.004) for the ultrasound model without age, 0.928 (SD 0.008) for the serum model, and 0.991 (SD 0.004) for the combined model. These estimates remain internal to the same cohort and do not address changes in case mix, operators, or equipment.

Bootstrap optimism-corrected AUCs were 0.798 for the age-only model, 0.983 for the age-adjusted ultrasound model, 0.982 for the ultrasound model without age, 0.928 for the serum model, and 0.987 for the combined model. Optimism-corrected calibration slopes were 1.008, 0.825, 0.903, 0.951, and 0.484, respectively. The corresponding optimism-corrected Brier scores were 0.167, 0.052, 0.056, 0.105, and 0.038. These internal-validation measures are summarized in Table 5. Bootstrap optimism-corrected calibration curves for Models 1–3 are shown in Figure 4.

Age-Stratified Performance

Because AWGS 2025 thresholds are age-specific and age is included in the primary model, we examined discrimination within the two AWGS age groups and applied the full-cohort Model 1 linear predictor within each stratum as a fixed-coefficient sensitivity analysis (Supplementary Table S2). Among participants aged 50-64 years ($n = 89$), only 6 sarcopenia events occurred; the within-stratum Model 1 AUC was 0.980 and the fixed-coefficient AUC was 0.916, so these values are reported only as unstable descriptive estimates. Among those aged 65 years or older ($n = 120$; 59 events), the age-only AUC was 0.603, the within-stratum Model 1 AUC was 0.986, and the fixed-coefficient Model 1 AUC was 0.973. Together with the full-cohort AUC of 0.984 for Model 1 without age, these results suggest that Model 1 discrimination cannot be explained solely by the age dependence of the AWGS definition. No interaction testing was performed.

Risk Reclassification and Discrimination Improvement

Categorical NRI at the 10% and 15% thresholds was 0.056 (95% CI, -0.011 to 0.123; bootstrap $P = 0.106$). Continuous NRI was 1.615 (bootstrap 95% CI, 1.431-1.782) and IDI was 0.085 (bootstrap 95% CI, 0.044-0.132). Thus, predicted probabilities were reordered, but the prespecified categorical thresholds did not show significant reclassification.

Clinical Utility: Decision Curve Analysis

Across threshold probabilities of 0.05-0.25, net benefit differed only modestly between the combined and ultrasound models. The combined-model values ranged from 0.292 to 0.306 and the ultrasound-model values from 0.276 to 0.299, with between-model differences of approximately 0.001-0.016. Because these thresholds were not linked to a prespecified management action, the decision-curve comparison is interpreted as exploratory (Figure 5).

Sensitivity Analyses

Adding sex, BMI, or both to the combined model produced little change in apparent discrimination (AUC 0.9956, 0.9962, and 0.9962, respectively, versus 0.9953 for the base combined model), but the maximum VIF rose from 5.04 to 5.95, 6.05, and 8.21, and the TyG_Z estimate increased in all three specifications. These analyses did not materially alter model-level discrimination but do not establish that sex or BMI are clinically unimportant (Supplementary Table S4).

Decile Patterns for EI and SII_Z

Restricted-cubic-spline analyses did not support significant nonlinearity for EI ($P = 0.107$) or SII_Z ($P = 0.261$). Observed sarcopenia proportions were generally higher in the upper EI deciles, whereas SII_Z was non-monotone at the upper end of its distribution (Supplementary Figs. S1 and S2). The plots

are descriptive and are not interpreted as evidence of a threshold, a dose-response relation, or a nonlinear effect.

Discussion

The primary finding was high apparent discrimination of the age-adjusted multiparametric rectus femoris ultrasound model for AWGS 2025-defined sarcopenia (AUC = 0.986; 95% CI: 0.976-0.997). Repeated cross-validation yielded a similar AUC, while the optimism-corrected calibration slope of 0.825 indicated some overfitting of predicted probabilities. Adding TyG_Z and SII_Z increased the apparent AUC by 0.0089 (DeLong P = 0.0447), but categorical reclassification was not significant and decision-curve gains were small; the combined-model calibration slope was 0.484. These results do not establish a clinically useful increment from the serum markers.

The age analyses address a separate concern: age is included in the model and also contributes to the AWGS 2025 diagnostic thresholds. Age alone had an AUC of 0.796, while Model 1 had an AUC of 0.986; removing age from Model 1 left the AUC at 0.984. Among participants aged 65 years or older, who accounted for 59 of the 65 events, the age-only AUC was 0.603 and the ultrasound-model AUC was 0.986. These findings suggest that the high discrimination of Model 1 cannot be explained solely by age. The younger stratum contained only six events and does not support a reliable subgroup performance estimate.

Sarcopenia was identified in 31.1% of the analytical cohort, higher than pooled estimates reported for broader Asian populations with T2DM 3,4. The difference is compatible with the hospitalized setting and the requirement to complete DXA and functional testing, so this proportion should not be interpreted as a population prevalence estimate. Participants with sarcopenia were older and had lower BMI, ASMI, and handgrip strength, slower gait speed, and poorer chair-stand performance. The imaging differences therefore occurred within a clinically coherent AWGS-defined phenotype rather than as isolated ultrasound abnormalities.

Several features may contribute to the high apparent discrimination, including separation of the ultrasound variables between groups, standardized acquisition, the hospitalized case mix, and development and evaluation in the same cohort. Morphologic ultrasound measures also overlap conceptually with the DXA-derived muscle-quantity component of the reference standard: rectus femoris thickness is a local size measure, whereas DXA ASMI is a whole-body appendicular measure. An exploratory model containing age, SWE_{mean}, and EI had an apparent AUC of 0.972, and the corresponding model without age had an AUC of 0.969. Substantial discrimination therefore remained when the size measure was omitted, although these estimates are subject to the same single-cohort setting.

The morphologic findings also differed by measurement. STD-RFthick and STD-RFc_{sa} were lower in participants with sarcopenia, whereas STD-RFc was similar between groups. Thickness and cross-sectional area more directly reflect local muscle size, while circumference can remain relatively preserved when non-contractile tissue occupies the fascial envelope. This pattern supports combining morphologic measures with tissue-characterization measures, while recognizing the conceptual overlap between ultrasound size measures and DXA-derived muscle quantity.

SWEmean was inversely associated with sarcopenia in the combined model, and resting stiffness was lower in participants with sarcopenia (9.84 vs. 14.36 kPa). Lower rectus femoris SWEmean has also been reported in COPD-related sarcopenia 22, but the biological basis is not resolved by this study. Passive muscle SWE is influenced by fiber architecture, muscle mass, fatty replacement, fibrosis, limb position, and low-level contraction, and these factors may act in different directions 9,13. Advanced glycation and fibrosis may increase extracellular-matrix stiffness, whereas atrophy, altered fiber organization, and myosteatosis may reduce the measured modulus of relaxed muscle. The present data establish an inverse cross-sectional association, not a specific histological mechanism.

Skeletal muscle is anisotropic, and SWE measurements are sensitive to fiber orientation, limb position, probe angle, muscle selection, and platform implementation 9-11. Transverse and longitudinal measurements are therefore not directly interchangeable. We standardized the transverse acquisition across participants because morphologic and elastographic measurements were obtained at the same site, but orientation-related variability remains. Absolute stiffness values should therefore be interpreted within this protocol and platform until cross-plane and cross-platform validation is available.

EI added information on grayscale tissue characteristics. Participants with sarcopenia had higher EI values, consistent with the greater echogenicity reported with intramuscular adipose and fibrous tissue, although their relative contributions cannot be separated without reference imaging or histology 8,21. In T2DM, myosteatosis has been linked to impaired insulin action, mitochondrial dysfunction, and poorer muscle performance 17,23. The concurrent findings of higher EI, lower SWEmean, lower ASMI, and poorer strength and physical performance are consistent with differences in both muscle quantity and tissue characteristics.

The laboratory findings describe the metabolic and inflammatory context of the observed muscle phenotype. The sarcopenia group had higher fasting glucose, triglycerides, TyG, SII, neutrophil, and platelet counts and a lower lymphocyte count, whereas HbA1c and diabetes duration were similar between groups. TyG and SII were associated with the concurrent outcome, but the cross-sectional design cannot determine whether they precede sarcopenia or reflect the inpatient metabolic and inflammatory state.

The combined model was less stable than the primary ultrasound model. Its conventional TyG_Z estimate was large and imprecise, and removal of the 27 observations flagged by influence diagnostics led to non-finite conventional estimates for TyG_Z, SWE_{mean}, and EI. Firth penalization produced finite coefficients and attenuated the TyG_Z estimate to 10.86 (95% CI: 2.59-88.33), but penalization does not correct calibration or transportability. The combined model and its individual coefficients are therefore presented as exploratory.

The apparent fit statistics and probability-ranking measures favored the combined model, but threshold-based results were less persuasive. Categorical NRI at 10% and 15% was not significant, and decision-curve net-benefit differences were small across the examined range. Continuous NRI and IDI describe changes in probability ranking rather than a clinical action. Because the thresholds were not linked to a prespecified management decision, these analyses should not be read as evidence of clinical benefit 24,25.

The incremental biomarker question is further constrained by the high discrimination of the base ultrasound model. With an apparent AUC of 0.986, little numerical room remained for a large AUC increase. This dataset cannot distinguish a genuinely small contribution from TyG and SII from a ceiling effect created by the base model. Their incremental role therefore remains unresolved.

Subgroup estimates are descriptive. The age strata contained 6 and 59 sarcopenia events, no interaction testing was performed, and the apparent AUC of 1.000 in participants younger than 65 years is too unstable for substantive interpretation. The age-stratified analysis is used here only to examine whether age alone explains the primary-model discrimination.

Multiparametric rectus femoris ultrasound may be useful as a candidate case-finding approach in hospitalized patients with T2DM, but it is not a substitute for DXA, BIA, handgrip testing, or the AWGS diagnostic assessment. Clinical implementation will require external validation and prospective evaluation of how an ultrasound-first strategy affects downstream testing and patient management.

The study combined the updated AWGS 2025 framework with blinded ultrasound acquisition, a reliability substudy based on repeat image acquisition, and multiparametric assessment using B-mode imaging, SWE, and quantitative EI. Model evaluation included discrimination, calibration, influence diagnostics, reclassification, and decision-curve analysis rather than AUC alone.

The main constraints are the cross-sectional, single-centre hospitalized design; the limited event count; and incomplete documentation of some clinical and study-conduct variables. Consecutive entry could not be verified, and the 209-patient analytical cohort may differ from the broader T2DM population in both selection and disease spectrum. The combined model was especially sensitive to influential

observations, and internal resampling cannot establish performance in a new population. Diabetes treatment, nutrition, physical activity, detailed diabetic complications, and comorbidity burden were not systematically available. Sex and BMI were examined only in sensitivity analyses. The reference standard used height-adjusted ASMI, so alternative muscle-mass adjustments were not assessed. SWE was acquired in the transverse plane, and both SWE and EI depend on acquisition conditions and system settings. Finally, the exact interval between ultrasound and the reference assessments, and blinding of DXA and functional assessors to ultrasound results, were not separately documented, although all assessments occurred during the same hospitalization. These issues define the scope of the present estimates and the priorities for external validation.

Prospective multicentre studies should validate the fixed model coefficients and calibration, determine whether protocol-specific ultrasound cutoffs reproduce across operators and systems, compare longitudinal and transverse SWE acquisition, and test ultrasound-first case-finding strategies against standard AWGS assessment in clinical workflows.

In conclusion, multiparametric rectus femoris ultrasound showed high apparent discrimination for AWGS 2025-defined sarcopenia in this hospitalized single-centre cohort, and the age analyses suggest that the finding cannot be explained solely by age. Adding TyG and SII produced only a small change in apparent discrimination; threshold-based reclassification was not significant, decision-curve gains were small, and the combined model showed marked coefficient and calibration instability. Because the base ultrasound model was already highly discriminative, the incremental role of the serum markers remains unresolved. Multiparametric rectus femoris ultrasound merits external validation as a candidate case-finding approach before clinical use.

References

1. Cruz-Jentoft, A. J. et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 48, 16-31; 10.1093/ageing/afy169 (2019).
2. Chen, L. K. et al. Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment. *J. Am. Med. Dir. Assoc.* 21, 300-307.e2; 10.1016/j.jamda.2019.12.012 (2020).

3. Yogesh, M., Patel, M., Gandhi, R., Patel, A. & Kidecha, K. N. Sarcopenia in type 2 diabetes mellitus among Asian populations: prevalence and risk factors based on AWGS-2019: a systematic review and meta-analysis. *BMC Endocr. Disord.* 25, 101; 10.1186/s12902-025-01935-y (2025).
4. Wu, Y. C., Mao, M. T., Huang, H. E., Huang, C. N. & Liao, W. C. Prevalence and risk factors of sarcopenia in Asian adults with type 2 diabetes: a systematic review and meta-analysis. *J. Diabetes Investig.* 17, 83–95 (2026).
5. Chen, L. K. et al. A focus shift from sarcopenia to muscle health in the Asian Working Group for Sarcopenia 2025 Consensus Update. *Nat. Aging* 5, 2164-2175; 10.1038/s43587-025-01004-y (2025).
6. Voulgaridou, G. et al. Diagnostic criteria and measurement techniques of sarcopenia: a critical evaluation of the up-to-date evidence. *Nutrients* 16, 436; 10.3390/nu16030436 (2024).
7. Rondaij, T., Rotovnik Kozjek, N., Jerele, C. & Jordan, T. Is there a place for ultrasound in diagnosing sarcopenia? *Radiol. Oncol.* 59, 153-167; 10.2478/raon-2025-0035 (2025).
8. Oranchuk, D. J., Bodkin, S. G., Boncella, K. L. & Harris-Love, M. O. Exploring the associations between skeletal muscle echogenicity and physical function in aging adults: a systematic review with meta-analyses. *J. Sport Health Sci.* 13, 820-840; 10.1016/j.jshs.2024.05.005 (2024).
9. Zhang, C. & Kang, L. Ultrasound elastography for the assessment of sarcopenia. *J. Clin. Med.* 15, 2566; 10.3390/jcm15072566 (2026).
10. Romano, A., Staber, D., Grimm, A., Kronlage, C. & Marquetand, J. Limitations of muscle ultrasound shear wave elastography for clinical routine: positioning and muscle selection. *Sensors* 21, 8490; 10.3390/s21248490 (2021).
11. Wang, X. et al. Assessment of ultrasound shear wave elastography within muscles using different region of interest sizes, manufacturers, probes and acquisition angles: an ex vivo study. *Quant. Imaging Med. Surg.* 12, 3227-3237; 10.21037/qims-21-1072 (2022).

12. Chen, M. et al. The value of ultrasound measurement of muscle thickness at different sites and shear wave elastography in Parkinson's disease with sarcopenia: a pilot study. *Front. Neurosci.* 17, 1254859; 10.3389/fnins.2023.1254859 (2023).
13. Galasso, L. et al. Echoes of muscle aging: the emerging role of shear wave elastography in sarcopenia diagnosis. *Diagnostics* 15, 2495; 10.3390/diagnostics15192495 (2025).
14. Stringer, H. J. & Wilson, D. The role of ultrasound as a diagnostic tool for sarcopenia. *J. Frailty Aging* 7, 258–261 (2018).
15. Simental-Mendía, L. E., Rodríguez-Morán, M. & Guerrero-Romero, F. The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects. *Metab. Syndr. Relat. Disord.* 6, 299-304; 10.1089/met.2008.0034 (2008).
16. Xu, M. et al. Associations between the triglyceride-glucose index and the odds of low muscle mass and sarcopenia: a systematic review and meta-analysis. *BMC Endocr. Disord.* 26, 118; 10.1186/s12902-026-02250-w (2026).
17. Ghosh, M. et al. Interplay between sarcopenia and type 2 diabetes: mechanisms, implications, and therapeutic prospects. *Nutr. Diabetes* 16, 18; 10.1038/s41387-026-00431-z (2026).
18. Hu, B. et al. Systemic immune-inflammation index predicts prognosis of patients after curative resection for hepatocellular carcinoma. *Clin. Cancer Res.* 20, 6212-6222; 10.1158/1078-0432.CCR-14-0442 (2014).
19. Xie, S. & Wu, Q. Association between the systemic immune-inflammation index and sarcopenia: a systematic review and meta-analysis. *J. Orthop. Surg. Res.* 19, 314; 10.1186/s13018-024-04808-7 (2024).
20. Chinese Diabetes Society. Guideline for the prevention and treatment of diabetes mellitus in China (2024 edition). *Chin. J. Diabetes Mellitus* 17, 16-139; 10.3760/cma.j.cn115791-20241203-00705 (2025).

21. Perkisas, S. et al. Application of ultrasound for muscle assessment in sarcopenia: 2020 SARCUS update. *Eur. Geriatr. Med.* 12, 45-59; 10.1007/s41999-020-00433-9 (2021).
22. Deng, M. et al. Ultrasonic elastography of the rectus femoris, a potential tool to predict sarcopenia in patients with chronic obstructive pulmonary disease. *Front. Physiol.* 12, 783421; 10.3389/fphys.2021.783421 (2022).
23. Prokopidis, K. & Duque, G. Association of intermuscular and intramuscular fat with insulin resistance, type 2 diabetes, and metabolic syndrome: a systematic review. *Clin. Nutr.* 64, 106727; 10.1016/j.clnu.2026.106727 (2026).
24. Vickers, A. J., van Calster, B. & Steyerberg, E. W. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn. Progn. Res.* 3, 18; 10.1186/s41512-019-0064-7 (2019).
25. Van Calster, B. et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Eur. Urol.* 74, 796-804; 10.1016/j.eururo.2018.08.038 (2018).

Acknowledgements

This study was supported by the Health Science and Technology Project of Anhui Provincial Health Commission (Grant No. AHWJ2025Ab40325). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization, J.M. and J.F.; Methodology, J.M., Y.H. and M.W.; Formal analysis, Y.H. and M.W.; Investigation, Y.H., D.X., Z.S. and S.G.; Data curation, Y.H., M.W., Z.S. and D.X.; Visualization, J.M.; Funding acquisition, M.W.; Writing—original draft, Y.H. and J.M.; Writing—review & editing, J.F., J.M. and M.W.; Supervision, J.F. and J.M.; Project administration, J.M.

All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Ethics declarations

Ethics approval and consent to participate: The study protocol was approved by the Institutional Ethics Committee of the Second People's Hospital of Wuhu (approval no. 2024-KY-231) and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before enrollment.

Consent for publication: Not applicable.

Figure Legends

Figure 1. Participant flow diagram. Of 297 patients with T2DM who underwent rectus femoris ultrasound, 88 were excluded because of incomplete DXA ($n = 42$), functional assessments ($n = 28$), or prespecified laboratory assessments ($n = 18$). The final analytical cohort included 209 participants, including 65 with sarcopenia (31.1%).

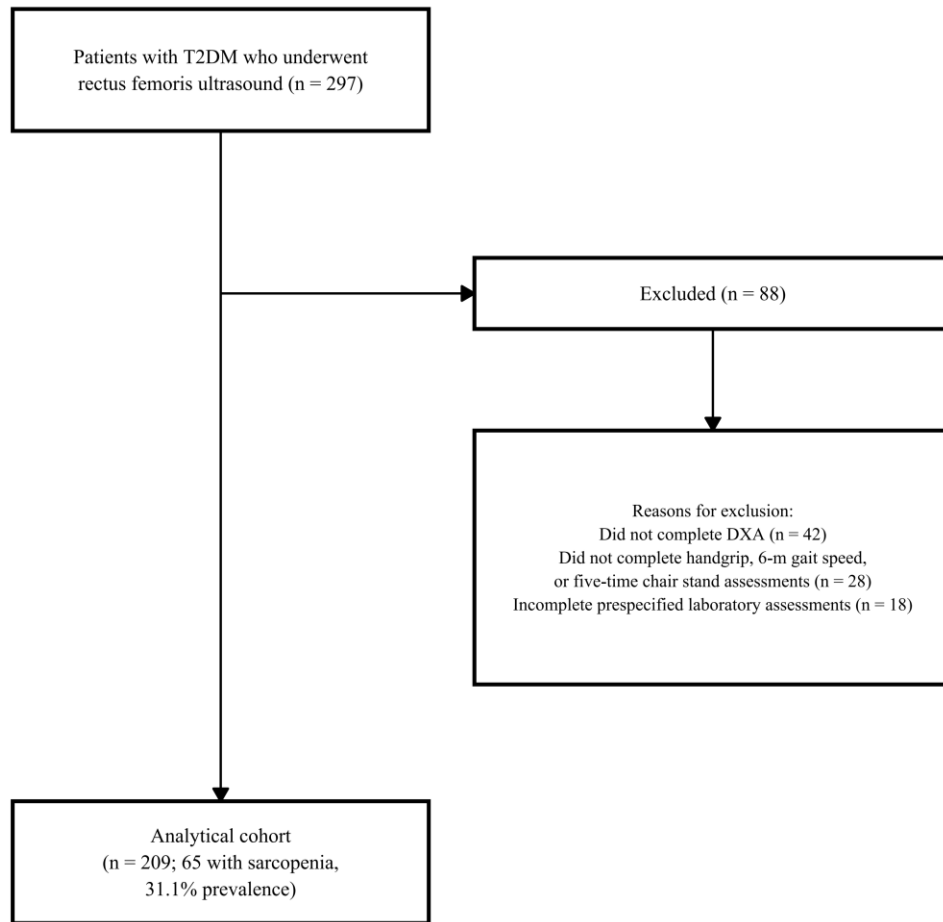
Figure 2. Forest plot of the conventional maximum-likelihood combined model (Model 3). Odds ratios and 95% Wald confidence intervals are shown on a logarithmic scale. Model 3 included age, STD-RFthick, SWEmean, EI, TyG_Z, and SII_Z. TyG_Z and SII_Z are expressed per 1-SD increase. These estimates are from conventional maximum-likelihood logistic regression; Firth penalized sensitivity estimates are reported separately in Supplementary Table S1.

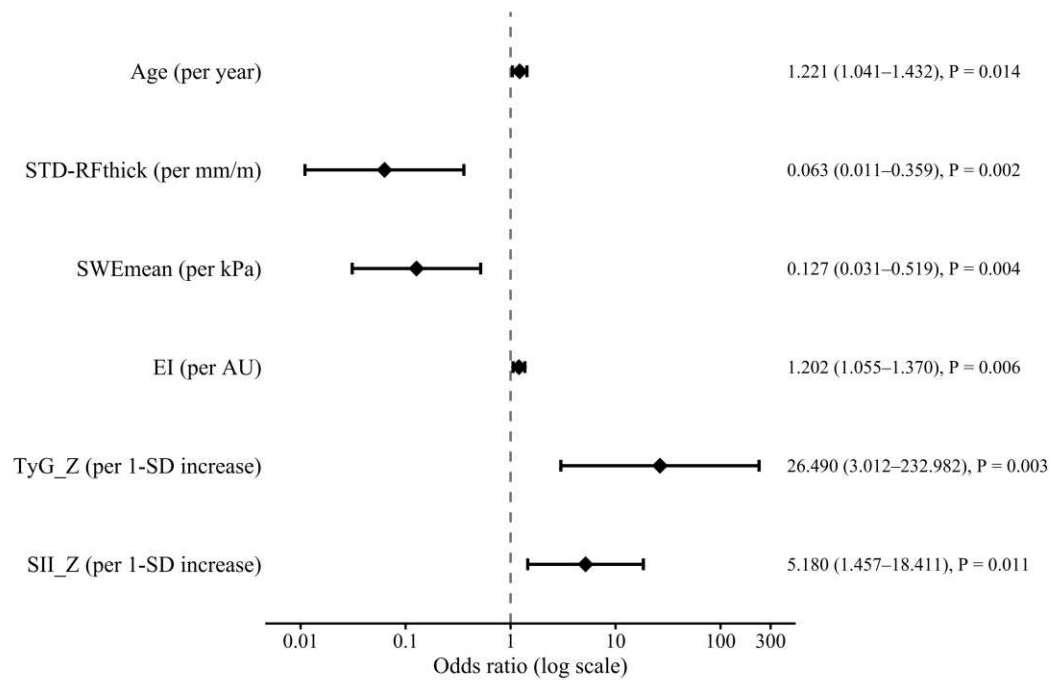
Figure 3. Receiver operating characteristic curves for the three prespecified logistic regression models. Model 1 included age, STD-RFthick, SWEmean, and EI; Model 2 included age, TyG_Z, and SII_Z; and Model 3 included all six predictors. Apparent AUCs in the development cohort were 0.986 (95% CI, 0.976–0.997), 0.932 (95% CI, 0.900–0.965), and 0.995 (95% CI, 0.990–1.000), respectively. Internal-validation results are reported separately in Table 5.

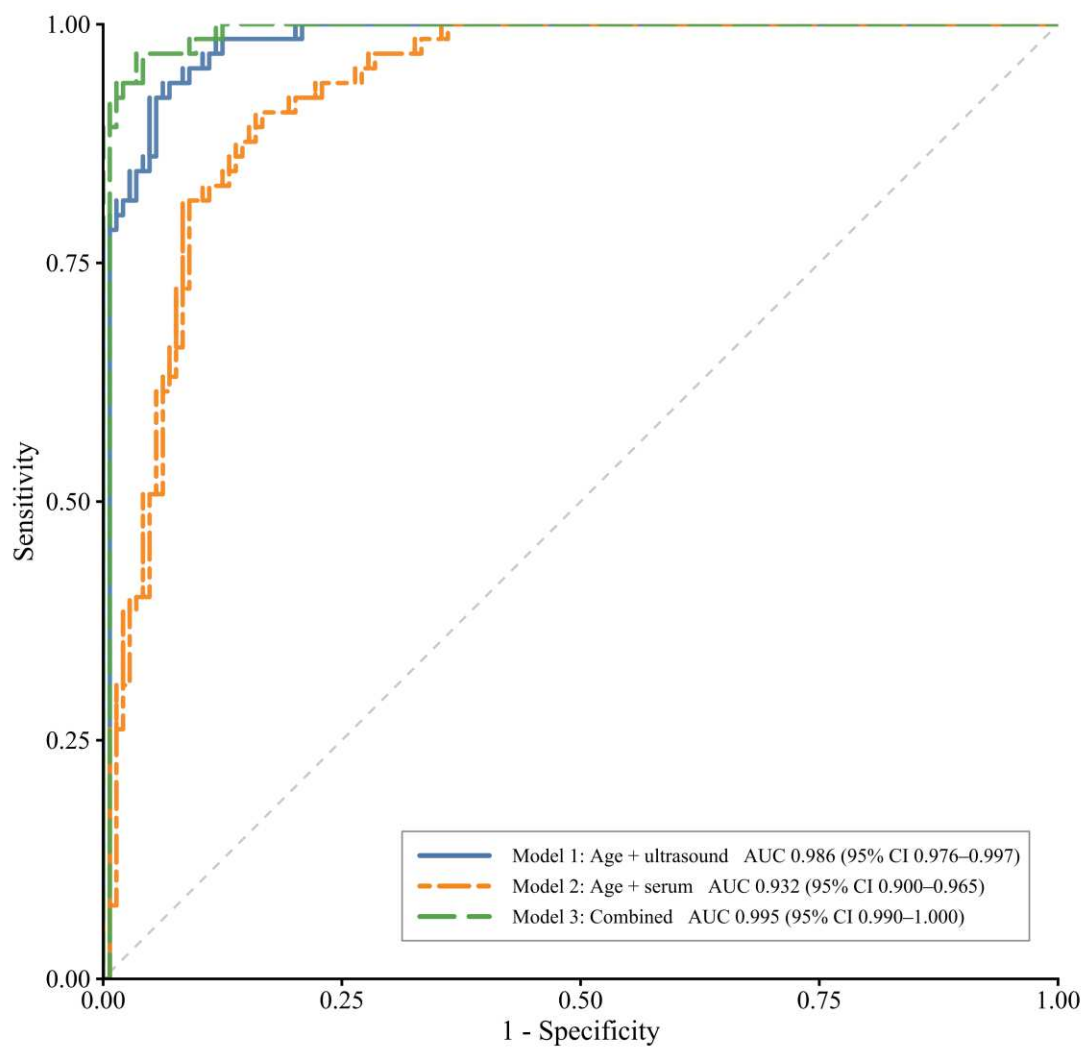
Figure 4. Bootstrap optimism-corrected calibration curves for the three logistic regression models. The diagonal dashed line represents ideal calibration. Model 1 included age, STD-RFthick, SWEmean,

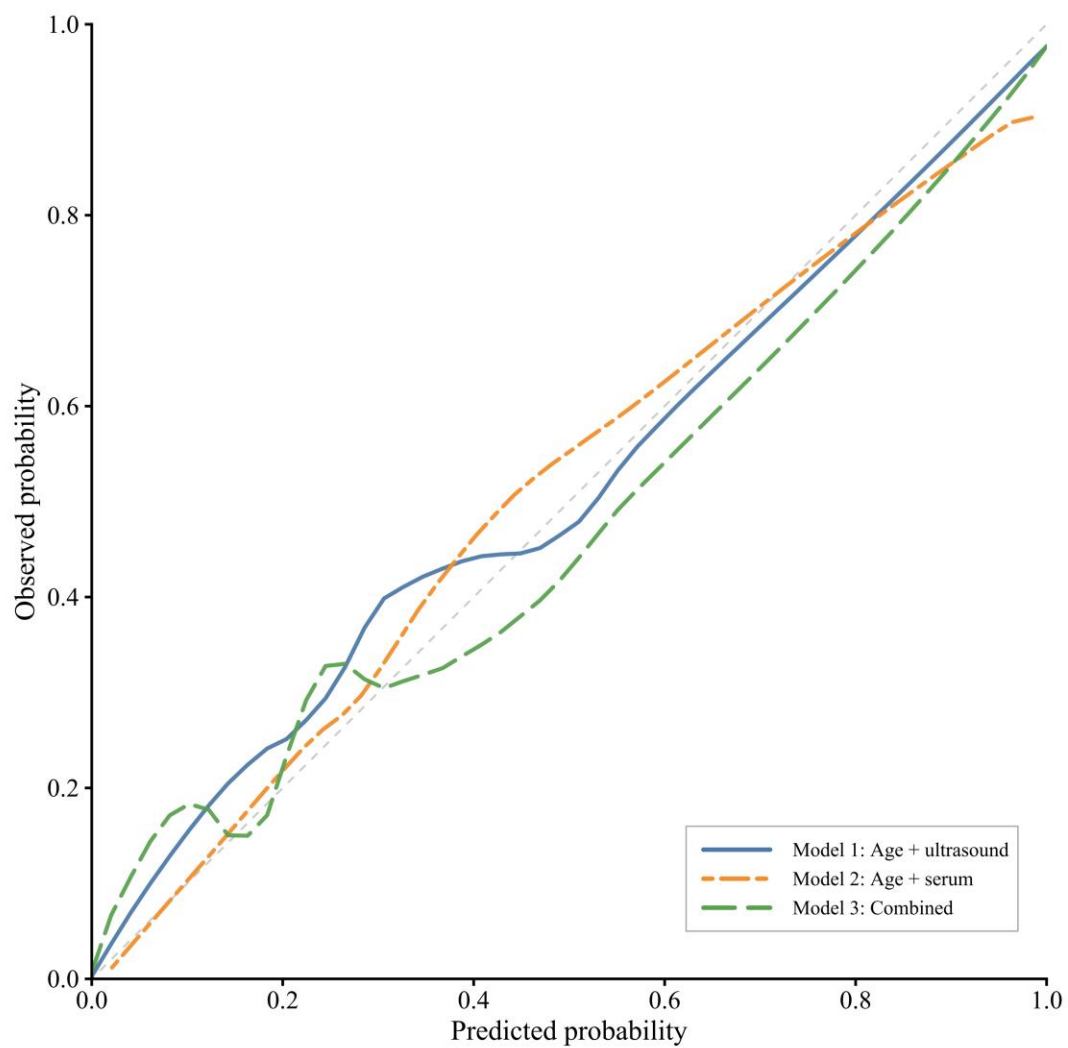
and EI; Model 2 included age, TyG_Z, and SII_Z; and Model 3 included all six predictors. Calibration curves were obtained using 1,000 bootstrap resamples. Optimism-corrected calibration slopes were 0.825, 0.951, and 0.484 for Models 1, 2, and 3, respectively.

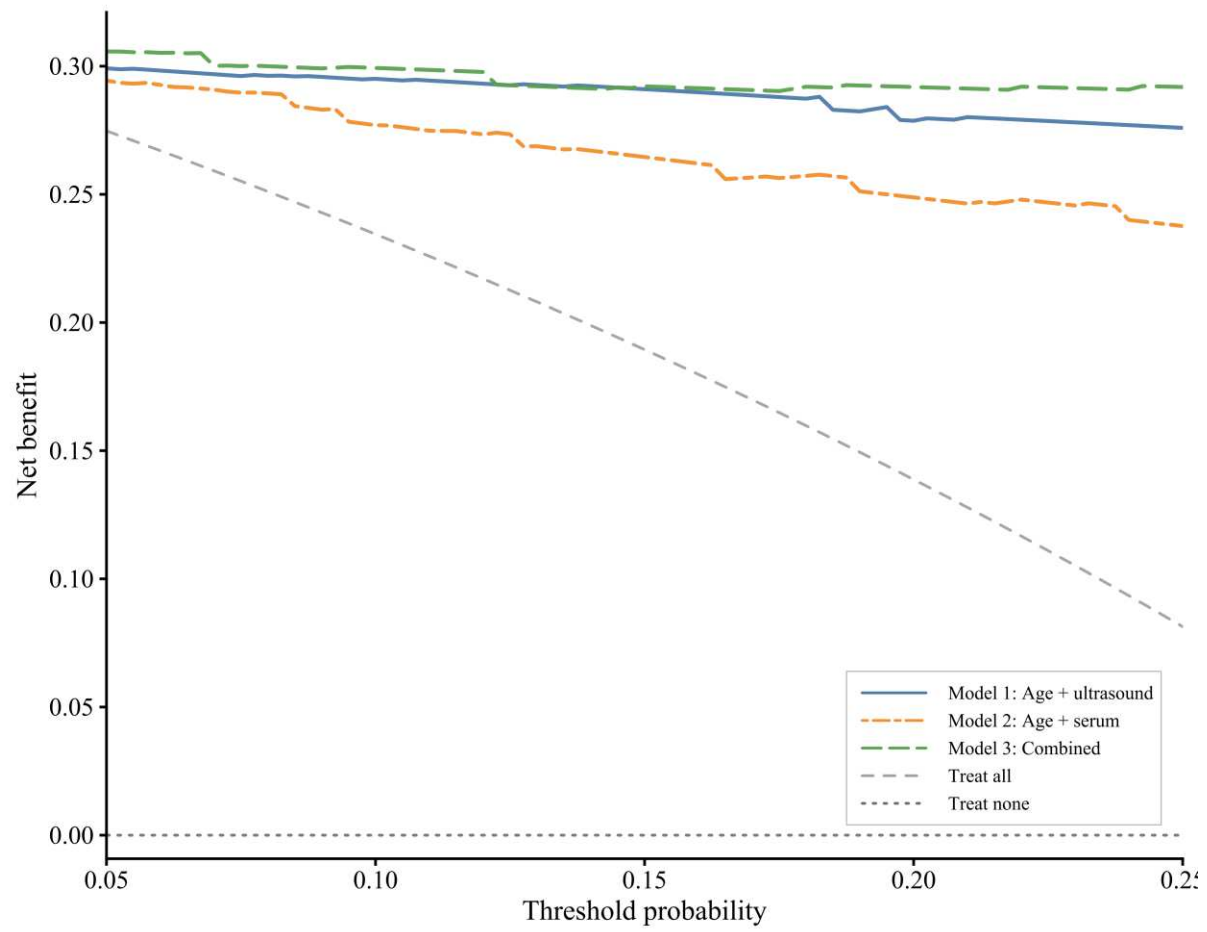
Figure 5. Decision-curve analysis for the three logistic regression models. Net benefit is shown across threshold probabilities from 0.05 to 0.25. Model 1 included age, STD-RFthick, SWEmean, and EI; Model 2 included age, TyG_Z, and SII_Z; and Model 3 included all six predictors. The “treat all” and “treat none” strategies are shown for reference. Differences in net benefit between the combined and ultrasound models were small across the examined thresholds; these thresholds were exploratory and were not linked to a prespecified clinical management action.











Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [05Tables.docx](#)
- [07SupplementaryMaterial.docx](#)