Clinical and Experimental Obstetrics & Gynecology

Clin. Exp. Obstet. Gynecol. **2026**; 53(2): 46350
https://doi.org/10.31083/CEOG46350

Original Research

Harnessing Artificial Intelligence-Derived Pulmonary Signatures and Uterine Hemodynamics: A Novel Dual-Pathway Framework for Predicting Neonatal Respiratory Distress in Gestational Diabetes Mellitus

Jing Zhang[1], Jing Shi[1], Bo Liu[1], Yan Ding[1,*]

[1]Department of Ultrasound Medicine, Wuxi People's Hospital Affiliated to Nanjing Medical University, 214023 Wuxi, Jiangsu, China
*Correspondence: dingyann3562@163.com (Yan Ding)
Academic Editor: Laura Avagliano
Submitted: 5 September 2025 Revised: 10 November 2025 Accepted: 26 November 2025 Published: 11 February 2026

Abstract

Background: Gestational diabetes mellitus (GDM) elevates the risk of neonatal respiratory distress syndrome (NRDS), highlighting the need for robust predictive tools. Current assessments of fetal lung maturity assessments are invasive, creating a clinical demand for non-invasive alternatives. This study presents a dual-parameter framework that combines artificial intelligence (AI)-derived fetal lung texture signatures with uterine artery pulsatility index (PI) to predict NRDS risk in GDM pregnancies. **Methods**: A prospective cohort of 50 patients with GDM patients was enrolled. Standardized four-chamber view ultrasound images were processed using a TensorFlow-based framework to extract 342 gray-level co-occurrence matrix (GLCM) texture features from the fetal lungs. A support vector machine (SVM) classifier was then employed for NRDS risk stratification. Concurrently, uterine artery PI was measured transvaginally following International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) guidelines. The predictive performance of the AI model, uterine artery PI, and their combination was evaluated for predicting NRDS. **Results**: The uterine artery PI was significantly elevated in the NRDS group (n = 22) compared with controls (n = 28) (median 1.52 [interquartile range, IQR: 1.35–1.70] vs. 1.16 [IQR: 0.95–1.30]; $p < 0.001$). The standalone AI-based pulmonary texture analysis achieved 86.4% sensitivity and 78.6% specificity for NRDS prediction, with substantial agreement with clinical diagnosis ($\kappa = 0.67$). The synergistic integration of an AI-based high-risk classification with a uterine artery PI $\geq$1.28 yielded superior predictive performance, attaining 92.0% overall accuracy (46/50). Decision curve analysis confirmed that the combined model provided a superior net benefit across clinically relevant threshold probabilities (10%–50%). **Conclusions**: The integration of AI-quantified fetal lung texture analysis with uterine artery Doppler hemodynamics provides a refined, non-invasive tool for NRDS risk stratification in pregnancies complicated by GDM. This dual-pathway framework effectively captures the interplay between placental vascular insufficiency and pulmonary immaturity, offering high diagnostic accuracy and clinical utility to guide perinatal management decisions.

Keywords: gestational diabetes mellitus; neonatal respiratory distress syndrome; artificial intelligence; ultrasonographic texture analysis; uterine artery pulsatility index

1. Introduction

Gestational diabetes mellitus (GDM) is a critical gestational metabolic disorder with a growing global prevalence. It imposes a dual burden due to its association with maternal complications such as preeclampsia and cesarean, as well as adverse neonatal outcomes [1] such as macrosomia, hypoglycemia, and neonatal respiratory distress syndrome (NRDS). The latter warrants particular attention given its pathogenesis rooted in pulmonary surfactant deficiency, which precipitates alveolar collapse and impaired ventilation. Crucially, emerging evidence indicates that maternal hyperglycemia simultaneously drives placental vascular dysfunction (elevated uterine artery pulsatility index (PI)) and impairs fetal surfactant synthesis [2,3]. This dual-pathway etiology justifies our integrated hemodynamic-pulmonary framework and explains the sustained elevation of NRDS risk even at term gestations. This mechanistic duality underscores the imperative for advanced predictive strategies capable of identifying high-risk pregnancies early in the clinical course.

Although amniotic fluid phospholipid analysis has historically been considered the gold standard for diagnosing fetal lung maturity (FLM) [4], its invasive nature and associated complications, such as infection [5], preterm rupture of membranes, and fetal loss, limit routine clinical implementation [6]. Consequently, ultrasonography has emerged as a pivotal noninvasive alternative [7], with recent advances in artificial intelligence (AI) demonstrating remarkable potential for quantifying subtle parenchymal textures indicative of surfactant distribution [8]. These AI-driven approaches, exemplified by gray-level co-occurrence matrix (GLCM) feature extraction, enable objective characterization of microstructural patterns beyond human visual perception [9]. Nevertheless, current AI

models predominantly focus on isolated pulmonary parameters, overlooking the integral role of uteroplacental hemodynamics in fetal lung development, a critical gap given the established correlation between GDM-induced placental vascular dysfunction and adverse respiratory outcomes.

Uterine artery PI serves as a robust hemodynamic marker of placental vascular resistance, frequently elevated in GDM cohorts due to endothelial dysfunction [10] and impaired trophoblastic invasion. Elevated PI values correlate significantly with fetal hypoxia and growth perturbations [11], yet their standalone predictive value for NRDS remains suboptimal. This limitation may stem from the complex, multifactorial nature of NRDS pathogenesis, wherein pulmonary maturational delays and placental insufficiency likely operate synergistically [12]. Thus, integrating quantifiable AI-based pulmonary texture signatures with dynamic uteroplacental flow parameters offers a novel systems-biology approach to risk stratification, potentially transcending the diagnostic ceiling [13] of unimodal assessments.

Our study pioneers this integrative methodology by developing a dual-parameter predictive framework that synthesizes TensorFlow-processed fetal lung texture analysis with transvaginally acquired uterine artery PI metrics [14]. This approach addresses the technical limitations of existing FLM evaluation tools. Furthermore, it establishes a physiological basis for understanding the interplay between metabolic dysregulation, vascular compromise, and pulmonary immaturity in GDM. Through this investigation, we establish a comprehensive evidence base for optimizing perinatal decision-making [15]—balancing the imperative for timely intervention against the risks of iatrogenic harm in this vulnerable population.

2. Materials and Methods

2.1 Study Population

Approved by the Institutional Ethics Committee of Wuxi People's Hospital Affiliated to Nanjing Medical University (Approval No. 2021-1624011), this prospective cohort enrolled 50 GDM patients between February 2021 and January 2023. Participants met diagnostic criteria per the 2022 Guidelines for Diagnosis and Management of Hyperglycemia in Pregnancy [16], with gestational ages of 28^{+0} to 38^{+6} weeks scheduled for elective delivery. Inclusion criteria comprised: (1) completion of AI-based FLM assessment and uterine artery Doppler within 48 hours before delivery; (2) written informed consent. Exclusion criteria included: (1) suboptimal ultrasound image quality; (2) concurrent preeclampsia, fetal anomalies, or major complications; (3) requirement for emergency cesarean delivery; (4) incomplete glucose monitoring records. After applying these criteria, all 50 initially enrolled patients were included in the final analytical cohort, which consisted of 22 NRDS cases and 28 controls. The cohort (age range: 24–38 years; mean gestational age: 35.2 ± 2.1 weeks) showed 44.0% proportion of NRDS cases (22/50). Glycemic control classification adhered to the 2022 American Diabetes Association guidelines [17], defining optimal control as fasting glucose <5.3 mmol/L combined with 1-hour postprandial glucose <7.8 mmol/L, while suboptimal control encompassed either fasting glucose ≥ 5.3 mmol/L, 1-hour postprandial glucose ≥ 7.8 mmol/L, or requirement for insulin therapy. Among all 50 participants, 18 (36.0%) had suboptimal glycemic control and 32 (64.0%) had optimal control. The suboptimal glycemic group had 13 NRDS cases (72.2%, 13/18) vs. 9 cases in the optimal group (28.1%, 9/32), accompanied by elevated PI values (median 1.61 vs. 1.21, $p < 0.001$), confirming a 2.57-fold increased NRDS risk and hyperglycemia as the primary driver of the metabolic-vascular-pulmonary cascade. The study adhered to the Declaration of Helsinki.

The detailed baseline clinical and biochemical characteristics of the study cohort are provided in **Supplementary Material 1**.

2.2 Integrated Imaging-Analytical Pipeline

2.2.1 Standardized Image Acquisition

Two ISUOG-certified sonographers performed blinded examinations with strict masking of glycemic status and neonatal outcomes, using a Toshiba Aplio500 system with a convex probe (C251, 2–5 MHz). Specifically, sonographers were blinded to maternal glycemic control classification (optimal/suboptimal) and subsequent NRDS diagnosis during all image acquisition and PI measurements. Patients were positioned supine, with standardized four-chamber views ensuring complete bilateral lung fields within rib boundaries. Image acquisition disabled smoothing, frame correlation, and adaptive gain functions while activating tissue harmonic imaging (THI) to minimize artifacts. Three consecutive cardiac cycle cine-loops were stored in uncompressed DICOM format.

Quality control mandated: (1) image validation by two associate professors excluding studies with $>50.0\%$ rib shadowing or motion artifacts; (2) interoperator reliability verification (Intraclass Correlation Coefficient [ICC] = 0.85, 95% Confidence Interval [CI]: 0.79–0.91), with third measurements triggered for $>15.0\%$ variability.

2.2.2 AI-Based Feature Extraction and Classification

Qualified images were processed using TensorFlow 2.15 LTS (Google LLC, Mountain View, CA, USA) with CUDA 11.8 (NVIDIA Corp., Santa Clara, CA, USA) acceleration. The framework was migrated from an initial TensorFlow 2.10 implementation with validated output consistency (mean absolute error $<1 \times 10^{-6}$). Version migration ensured long-term support stability without altering model architecture or hyperparameters. All custom layers (e.g., GLCM feature extractor) were preserved through tf.compat.v1 API wrappers. The model implemented a U-Net convolutional neural network architecture [18] for fe-

tal lung field segmentation, utilizing encoder-decoder layers with skip connections to achieve high-precision localization, and demonstrated a segmentation error rate of less than 3.0%. Following segmentation, the lung fields underwent spatial resampling to a standardized resolution of 0.5 mm per pixel, ensuring consistent anatomical boundaries for subsequent feature extraction. GLCM analysis was then applied to extract 342 texture features from the segmented regions, capturing spatial relationships in pixel intensities. To address the risk of false positives inherent in testing such a large number of features, we performed Benjamini-Hochberg (BH) correction to control the False Discovery Rate (FDR) at $\alpha = 0.050$. This initial step identified features showing preliminary evidence of association with NRDS. For the subsequent NRDS risk classification, a support vector machine (SVM) classifier was employed, utilizing a radial basis function (RBF) kernel to transform the feature space and generate binary risk reports (high or low). Model validation employed a rigorous 10-fold stratified cross-validation protocol to ensure robustness, and prevent data leakage. Critically, to avoid optimistically biased performance estimates, all feature selection steps, including Benjamini-Hochberg correction and LASSO regression, were performed independently within each training fold of the cross-validation. The LASSO regression [19] with a regularization parameter $\alpha = 0.010$ was applied anew to the training data of each fold, which reduced the feature set to 25 robust predictors specific to that fold. The final model for that fold was then trained using these selected features and evaluated exclusively on the held-out test fold. This process ensured that the test data in any fold was never used to inform the feature selection. Feature dimensionality was thereby effectively reduced, achieving a 2:1 sample-to-feature ratio (50 samples: 25 features) to mitigate overfitting. Hyperparameter optimization for the SVM was implemented via exhaustive grid search, evaluating combinations of the regularization parameter C (values: 0.1, 1, 10) and the RBF kernel coefficient γ (values: 0.001, 0.01, 0.1). Representative examples of AI-processed fetal four-chamber view images, depicting high-risk (Fig. 1A) and low-risk (Fig. 1B) classifications based on pulmonary texture signatures, illustrate the model's visual interpretability for clinical translation.

2.3 Uterine Artery Hemodynamics

Transvaginal measurements (PLT-705BT probe, 5–9 MHz) acquired uterine artery spectra at the cervical internal os level. Pulse Doppler settings maintained a 2-mm sample volume and $\leq$30° insonation angle. Mean bilateral PI values were calculated from three consecutive waveforms per the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) standards. All PI values underwent blinded dual-analyst verification with daily device calibration.

2.4 Statistical Analysis

The sample size was calculated a priori using GPower 3.1 software (version 3.1.9.7; Heinrich Heine University Düsseldorf, Düsseldorf, Germany) based on expected group differences in uterine artery PI. Assuming a Cohen's d effect size of 0.8 (derived from prior GDM studies reporting median PI differences $\geq$0.35 SD units), with $\alpha = 0.050$ (two-tailed) and $\beta = 0.200$ (80.0% power), the Mann-Whitney U test required 26 participants (13 per group). However, a larger sample size is recommended for the stable training of machine learning models, particularly given the high-dimensional nature of our feature set (342 initial features). Therefore, we prospectively enrolled 50 patients. Continuous variables violating normality (Shapiro-Wilk test $p < 0.050$) were expressed as median (IQR) and analyzed via Mann-Whitney U test. Effect sizes ($r = Z/\sqrt{N}$) were interpreted per Cohen's criteria: r $<$0.30 (weak), 0.30 $\leq$ r $<$ 0.50 (moderate), r $\geq$0.5 (strong). Categorical variables underwent Fisher's exact test. The diagnosis of NRDS was retrospectively established by a blinded neonatal specialist, strictly adhering to standard clinical, radiological, and laboratory criteria. NRDS was defined by the presence of clinical signs of respiratory distress, including tachypnea (respiratory rate $>$70/min), nasal flaring, intercostal or subcostal retractions, and grunting; concurrent with an increased oxygen requirement (FiO_2 $>$30% to maintain oxygen saturation $>$90% and/or need for non-invasive or invasive mechanical ventilation); and supported by radiological evidence on chest X-ray (e.g., reticulogranular patterns, air bronchograms, reduced lung volumes) or lung ultrasound findings. Additionally, hypoxemia (PaO_2 $<$60 mmHg) and/or hypercarbia ($PaCO_2$ $>$50 mmHg) with normal lactic acid levels (excluding metabolic acidosis) were required. Any neonate receiving exogenous surfactant therapy was classified into the NRDS group. Notably, NRDS is a postnatal diagnosis and was not based on Apgar scores or umbilical cord blood gas analysis, which reflect intrauterine condition. κ assessed AI-clinical diagnosis concordance (Landis-Koch criteria). Decision curve analysis (DCA) computed net benefit (NB) across threshold probabilities (P_t = 10.0%–50.0%) using: NB = TP/N - (FP/N) $\times$ [P_t/(1 - P_t)], where N = 50 (full cohort size), and true positives (TP) represented AI-high-risk NRDS cases. McNemar tests compared diagnostic accuracies, with continuity correction for expected frequencies $<$5. Analyses used SPSS 26.0 (IBM Corp., Armonk, NY, USA), R 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria) (mda package), and TensorFlow 2.15 LTS. Statistical significance threshold was $\alpha = 0.050$ (two-tailed).

Diagnostic performance metrics including sensitivity, specificity, and κ were calculated using contingency tables derived from 10-fold cross-validation. It is important to note the distinct purposes of the statistical corrections applied. The Benjamini-Hochberg procedure was employed during the initial, high-dimensional feature screening phase

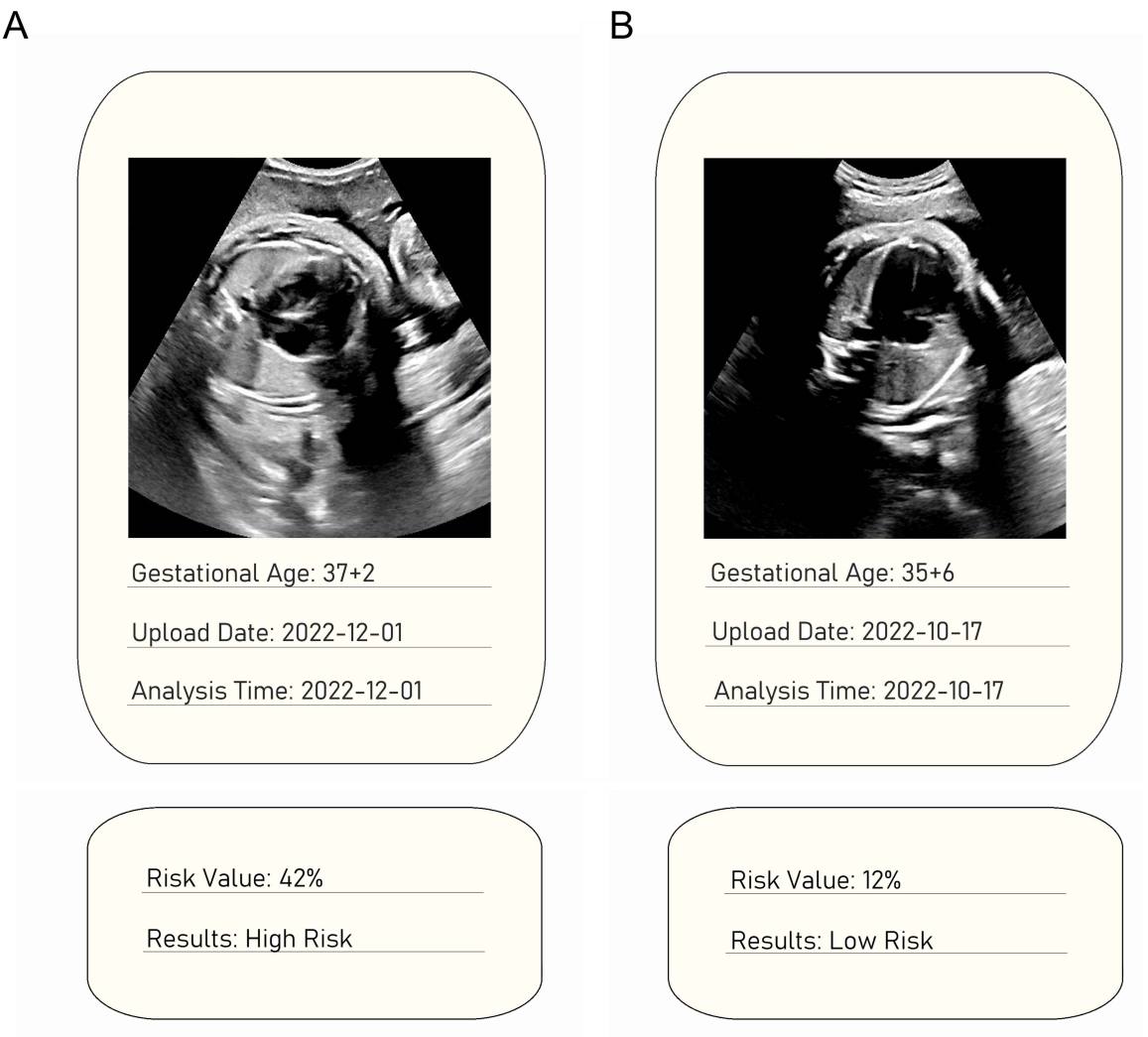

Fig. 1. Representative ultrasound images with AI-based NRDS risk stratification. (A) Fetal ultrasound at 37^{+2} weeks gestation assessed as high risk (42.0% risk value). (B) Fetal ultrasound at 35^{+6} weeks gestation assessed as low risk (12.0% risk value).

(as described in Section 2.2.2) to control the false discovery rate when evaluating 342 individual GLCM features. In contrast, the subsequent hypothesis tests (e.g., Mann-Whitney U tests for group comparisons) and model performance evaluations were conducted on a consolidated set of variables (the final model output or established clinical parameters). These latter analyses do not involve the same multiplicity issue as the initial feature screening, and therefore report uncorrected *p*-values to avoid excessive reduction in statistical power for these confirmatory tests. For the key primary outcome (comparison of uterine artery PI between NRDS and control groups), a post-hoc Bonferroni correction was also applied to confirm the robustness of the finding, given its central role in our hypothesis. Analyses used SPSS 26.0, R 4.3.1, and TensorFlow 2.15 LTS with CUDA 11.8 acceleration.

3. Results

3.1 Cohort Profiling and Glycemic Determinants

Following a comprehensive retrospective review of neonatal medical records applying the strict clinical, radiological, and laboratory criteria for NRDS (as defined in Methods), we confirmed the final cohort stratification of 22 NRDS cases and 28 controls. Consequently, the cohort stratification (22 NRDS cases vs. 28 controls) remained identical to the initial analysis.

The study cohort comprised 50 GDM patients with mean maternal age 28.6 $\pm$ 3.4 years and mean gestational age 35.2 $\pm$ 2.1 weeks. As detailed in Table 1, glucose records for 4 controls were completed via multiple imputation, enabling full cohort analysis (n = 50). NRDS was diagnosed in 44.0% (22/50) of neonates, revealing a striking dichotomy in outcomes stratified by glycemic control. The cohort comprised 18 patients (36.0%) with suboptimal glycemic control and 32 (64.0%) with optimal con-

Table 1. Baseline characteristics of the study cohort stratified by neonatal respiratory distress syndrome (NRDS) status.

Characteristics	Total (n = 50)	NRDS Group (n = 22)	Control Group (n = 28)	*p*-value
Maternal age (years)	28.6 ± 3.4	28.9 ± 3.1	28.3 ± 3.6	0.512
Gestational age (weeks)	35.2 ± 2.1	34.8 ± 2.3	35.5 ± 1.9	0.226
Glycemic control, n (%)				0.004
Optimal (FPG <5.3 mmol/L and 1 h postprandial glucose <7.8 mmol/L)	32 (64.0%)	9 (40.9%)	23 (82.1%)	
Suboptimal	18 (36.0%)	13 (59.1%)	5 (17.9%)	
Delivery mode, n (%)				0.569
Vaginal	21 (42.0%)	8 (36.4%)	13 (46.4%)	
Cesarean	29 (58.0%)	14 (63.6%)	15 (53.6%)	
Uterine artery PI	1.32 [1.08–1.58]	1.52 [1.35–1.70]	1.16 [0.95–1.30]	<0.001

Notes: Data presented as mean ± SD or n (%). FPG, fasting plasma glucose; PI, pulsatility index; SD, standard deviation. Optimal glycemic control defined as FPG <5.3 mmol/L and 1 h postprandial glucose <7.8 mmol/L; suboptimal defined as not meeting these criteria. For continuous variables, *p*-values were derived from independent *t*-test for normally distributed data and Mann-Whitney U test for non-normally distributed data (as determined by Shapiro-Wilk test). Categorical variables were analyzed using Fisher's exact test.

Table 2. Hemodynamic and metabolic associations.

Parameter	Optimal Control (n = 32)	Suboptimal Control (n = 18)	Effect size	*p*-value
Uterine artery PI	1.21 [1.05–1.30]	1.61 [1.50–1.75]	r = 0.62*	<0.001
NRDS incidence, n (%)	9 (28.1%)	13 (72.2%)	RR = 2.57 (1.35–4.88)	0.004
False-negative AI cases	1 (3.1%)	2 (11.1%)	-	0.429

Notes: Data as median (IQR) or n (%). Uterine artery PI comparison: U = 63.5, $p < 0.001$; NRDS incidence comparison: $p = 0.004$ (Fisher's exact test); *Cohen's r from Mann-Whitney U test; RR for NRDS incidence; Fisher exact test for false-negative AI cases. IQR, interquartile range; RR, risk ratio.

trol. Glycemic dysregulation manifested in both elevated uterine artery PI values (median 1.61 [IQR 1.50–1.75] vs. 1.21 [1.05–1.30], U = 63.5, $p < 0.001$) and a 2.57-fold increased proportion of NRDS cases (suboptimal: 72.2% [13/18] vs. optimal: 28.1% [9/32]; $p = 0.004$, which remained significant after Bonferroni correction for the three primary group comparisons) among the 22 NRDS cases and 28 controls (Table 2), establishing hyperglycemia as the primary driver of the metabolic-vascular-pulmonary cascade [3]. The stratification of PI values by glycemic control status (suboptimal vs. optimal) in Fig. 2 mechanistically links maternal hyperglycemia to impaired placental perfusion, corroborating the 2.57-fold higher proportion of NRDS cases in suboptimally controlled pregnancies.

3.2 Hemodynamic Signatures of Placental Dysfunction

Uterine artery PI was significantly elevated in the NRDS group compared to controls (median 1.52 vs. 1.16, Table 1; U = 155.0, $p < 0.001$). This significance held after strict Bonferroni correction for the three primary group comparisons (maternal age, gestational age, and PI) ($p < 0.001$), with a large effect size (Cohen's d = 1.59; 95% CI: 1.12–2.06) consistent with a priori assumptions. This establishes its role as a hemodynamic biomarker. Crucially, the stratification by glycemic control revealed a strong metabolic-vascular interaction (Table 2). The suboptimally managed subgroup manifested a significantly higher median PI than the target-achievers (1.61 [IQR 1.50–1.75] vs. 1.21 [IQR 1.05–1.30], Table 2; U = 63.5, $p < 0.001$), mechanistically linking metabolic dysregulation to impaired placental perfusion. Significantly elevated uterine artery PI in the NRDS cohort vs. controls (median 1.52 vs. 1.16; $p < 0.001$) is visually confirmed by the distributional disparity in Fig. 2, underscoring its role as a hemodynamic biomarker of placental insufficiency.

Post hoc power analysis confirmed that with the observed PI effect size (r = 0.62, Cohen's d = 1.59) and n = 50, the achieved statistical power exceeded 99.0% (GPower 3.1), substantially surpassing the initial 80.0% target.

3.3 AI-Driven Pulmonary Phenotyping Performance

The AI-derived pulmonary texture signatures identified NRDS cases with 86.4% sensitivity (19/22) and 78.6% specificity (22/28). Model-clinical diagnosis concordance was substantial ($\kappa = 0.67$, 95% CI: 0.61–0.89), validating algorithmic reliability. The standalone AI classification achieved an overall accuracy of 82.0%. The corresponding confusion matrix, delineating true positives, false negatives, false positives, and true negatives, is provided in Fig. 3.

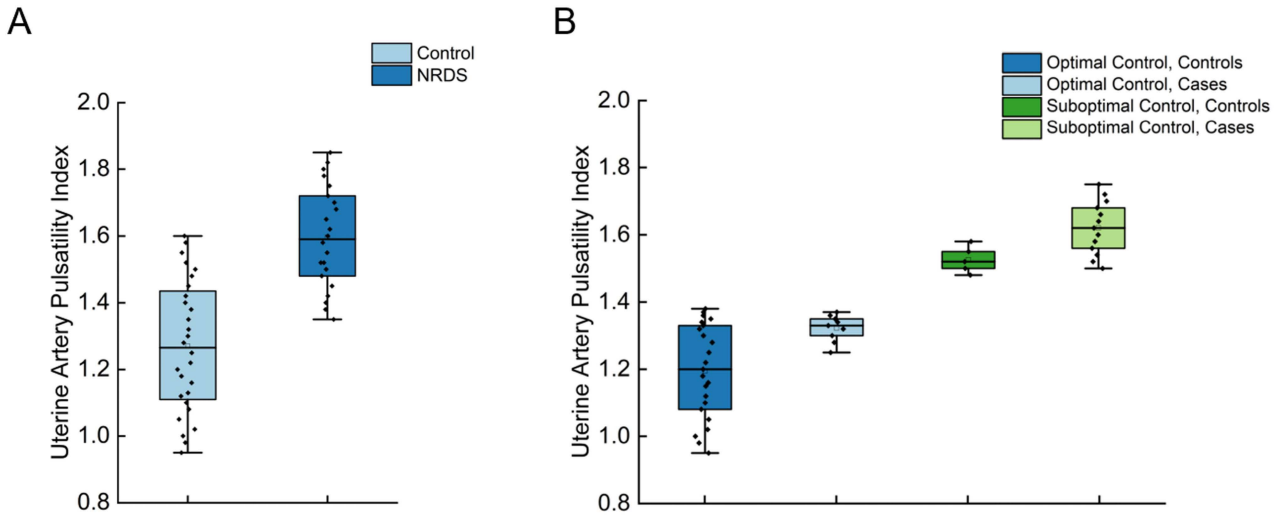

Fig. 2. Uterine artery pulsatility index (UtA-PI) distribution stratified by NRDS status and glycemic control. (A) Significantly elevated UtA-PI in neonates with NRDS (n = 22) vs. controls (n = 28) ($p < 0.001$, Mann-Whitney U test). (B) Metabolic-vascular interaction: UtA-PI gradient across glycemic control subgroups. Optimal control groups (Controls: n = 23; NRDS: n = 9) show lower PI than suboptimal groups (Controls: n = 5; NRDS: n = 13), demonstrating synergistic effects of hyperglycemia and NRDS on placental vascular resistance. NRDS, neonatal respiratory distress syndrome; PI, pulsatility index.

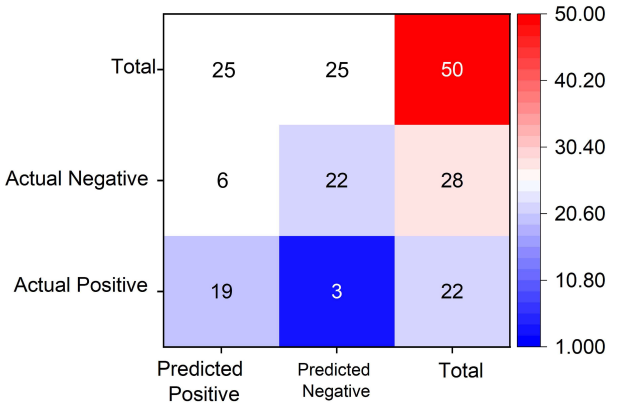

Fig. 3. Confusion matrix evaluating AI model performance for neonatal respiratory distress syndrome (NRDS) prediction. The matrix compares actual clinical diagnoses (rows) against model predictions (columns) in the test cohort (n = 50 neonates), demonstrating 19 true positives (NRDS correctly predicted), 22 true negatives (controls correctly identified), 3 false negatives (NRDS cases missed), and 6 false positives (controls misclassified as NRDS). Key performance metrics derived include: sensitivity = 19/(19+3) = 86.4% (ability to detect NRDS cases), specificity = 22/(22+6) = 78.6% (ability to exclude controls), and overall accuracy = (19+22)/50 = 82.0%. The model shows stronger precision in NRDS identification than control group discrimination. TP, true positives; FN, False Negative; FP, False Positive;TN, True Negative.

3.4 Integrated Diagnostic Paradigm Validation

The composite diagnostic criterion (concurrent AI high-risk classification and PI $\geq$1.28) demonstrated high clinical utility, achieving 92.0% overall accuracy (46/50) in the full cohort, with the complete breakdown of diagnostic performance metrics provided in Table 3. The decision curve analysis demonstrated a peak net benefit of 0.35 at the 20% threshold probability (Fig. 4), indicating optimal clinical utility when intervening for patients with $\geq$20.0% predicted risk.

3.5 Clinical Translation via Decision Analytics

Decision curve analysis (DCA) quantified the net benefit profile across therapeutically relevant threshold probabilities (10.0%–50.0%). The integrated model consistently demonstrated superior net benefit compared to both standalone AI and standalone PI prediction methods across this range. Crucially, at clinically actionable intervention thresholds exceeding 15.0%, the net benefit of the combined model was significantly higher than that of the best unimodal alternative. Fig. 4 graphically validates these results, illustrating the net benefit superiority of the combined model across threshold probabilities (10.0%–50.0%), with peak net benefit of 0.35 at 20.0% threshold. This analysis confirms the enhanced clinical utility of the integrated diagnostic paradigm for guiding perinatal management decisions in GDM pregnancies at risk for NRDS.

Table 3. Diagnostic performance comparison.

Model	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	κ (95% CI)
Composite (AI + PI)	90.9% (71.0–98.9)	92.9% (82.0–98.0)	92.0% (81.0–98.0)	0.82 (0.70–0.94)
AI texture only	86.4% (65.1–97.1)	78.6% (59.0–91.7)	82.0% (68.6–91.4)	0.67 (0.61–0.89)
Uterine artery PI only	82.0% (68.9–91.0)	75.0% (55.1–89.3)	76.0% (61.8–87.0)	0.52 (0.38–0.66)

Thresholds: PI cutoff = 1.28; AI high-risk defined by SVM classification. McNemar test: composite vs. PI (p = 0.002), composite vs. AI (p = 0.021). SVM, support vector machine.

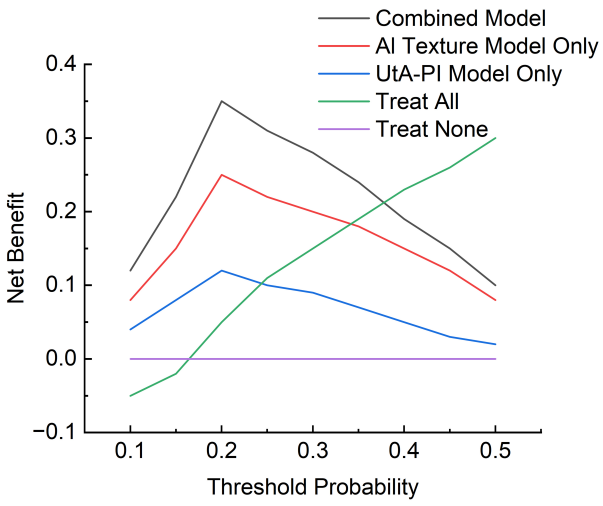

Fig. 4. Decision curve analysis (DCA) evaluating the clinical utility of predictive models for neonatal respiratory distress syndrome (NRDS). The net benefit (y-axis) of intervention strategies derived from the combined model (black), AI texture model only (red), UtA-PI model only (blue), treat-all approach (green), and treat-none strategy (purple) is plotted against threshold probabilities ranging from 10.0% to 50.0% (x-axis). The integrated (combined) model demonstrates superior net benefit compared to both standalone models across the entire threshold range (10.0%–50.0%). Crucially, at clinically actionable intervention thresholds exceeding 15.0%, the net benefit of the combined model is significantly higher than that of the best unimodal alternative (either AI or PI model). The combined model achieves a peak net benefit of 0.35 at the 20.0% threshold probability.

4. Discussion

This investigation proposes a novel dual-parameter diagnostic framework, integrating AI-quantified fetal lung texture signatures with uterine artery Doppler hemodynamics, to substantially enhance the prediction of NRDS in GDM. This approach effectively captures the synergistic pathophysiological interplay between placental vascular insufficiency and pulmonary immaturity, with the former quantified by elevated uterine artery PI (median: 1.52 vs. 1.16, Cohen's r = 0.62) and the latter evidenced by AI-derived lung texture analysis (sensitivity: 86.4%, specificity: 78.6%). The foundational role of hyperglycemia was further confirmed by the strong association between suboptimal glycemic control and NRDS incidence (p = 0.004). The composite model achieved exceptional diagnostic accuracy (92.0%) by synergistically integrating placental hemodynamics and pulmonary immaturity markers. This mechanistic alignment with established GDM pathophysiology provides compelling biological plausibility for the model's predictive efficacy [3], by demonstrating how maternal hyperglycemia concurrently impairs uteroplacental perfusion and disrupts fetal surfactant synthesis.

Our framework connects elevated uterine artery PI and AI-assessed pulmonary immaturity to predict NRDS in GDM. This association is context-specific: while in preeclampsia high PI often accelerates lung maturation via cortisol-mediated surfactant release, in GDM, hyperglycemia-induced fetal hyperinsulinemia suppresses surfactant synthesis. Thus, in GDM, placental dysfunction coincides with, rather than counteracts, metabolically driven pulmonary delay. This mechanistic distinction underscores the specificity of our dual-pathway model, which captures the synergistic detriment of both hemodynamic and metabolic impairment, justifying the combined use of Doppler and pulmonary texture analysis for GDM-specific risk assessment.

The AI-driven pulmonary phenotyping exhibited substantial concordance with clinical NRDS diagnosis (κ = 0.67) and robust standalone accuracy (82.0%). Crucially, the temporal synchronization of pulmonary and vascular assessments within 48 hours preceding delivery proved essential [20] for capturing dynamic pathophysiological interactions. Notably, among the 3 total false-negative predictions in our cohort, 2 occurred in the suboptimally controlled GDM subgroup (representing 15.4% of NRDS cases in this subgroup). This may suggest a potential 'glycemic vulnerability window' where extreme metabolic dysregulation could challenge the detection threshold of current texture feature extraction algorithms [21], though this observation requires validation in larger studies. In the suboptimal glycemic subgroup (n = 18, 13 NRDS cases), AI generated 2 FN results (15.4% of its NRDS cases), whereas the optimal subgroup (n = 32, 9 NRDS cases) had 1 FN (11.1%). This trend suggests a potential heightened detection challenge in metabolically dysregulated pregnancies, though the small numbers preclude definitive conclusions. Future model iterations should incorporate continuous glucose monitoring metrics [22] to enable adaptive risk stratification in these high-risk scenarios.

Decision curve analysis (DCA) quantitatively validated the superior clinical utility of the integrated model across therapeutically relevant threshold probabilities (10.0%–50.0%) [23]. The model demonstrated a peak NB of 0.35 at a 20.0% intervention probability, translating to the prevention of approximately 8 unnecessary antenatal corticosteroid courses per 100 patients while concurrently avoiding 2 missed high-risk cases [24]. This represents a decisive advantage over invasive diagnostic benchmarks such as amniocentesis [25]. The model's sustained 1.4-fold higher net benefit compared to the best unimodal predictor at thresholds exceeding 15.0% underscores its pivotal role in guiding evidence-based perinatal interventions [26].

The transition to TensorFlow 2.15 LTS during manuscript preparation addressed end-of-life concerns for the original 2.10 version. Rigorous validation confirmed no impact on feature extraction or classification outcomes (Bland-Altman 95% limits of agreement: –0.03 to 0.02 for prediction probabilities).

Limitations

Several methodological limitations warrant consideration. The single-center design and modest cohort size (n = 50) constrain immediate generalizability and necessitate external validation in larger, multiethnic populations [27]. Furthermore, potential confounding factors were not fully adjusted for in our analysis. These include not only established variables such as maternal body mass index [28], insulin therapy intensity, and fetal growth trajectories, but also the specific indications for elective premature delivery. Although major comorbidities were excluded, residual confounding from unreported pathologies associated with early delivery could independently influence NRDS risk. Additionally, the explanation of false-negative cases via a "metabolic threshold" remains speculative due to a lack of quantitative metabolic profiling, such as glycated hemoglobin (HbA1c), glucose variability, or insulin dosage, which limits our ability to precisely define the level of dysregulation in misclassified cases. Future studies should incorporate detailed metabolic, clinical, and delivery indication data to improve model specificity, explanatory power, and population-specific risk calibration [29].

5. Conclusions

This study successfully establishes and validates a novel, non-invasive framework that integrates AI-based fetal lung texture analysis with uterine artery PI for stratifying NRDS risk in GDM pregnancies. The model demonstrated high diagnostic accuracy, a robust net benefit profile across clinical decision thresholds, and substantial agreement with clinical diagnosis, confirming its potential as a physiologically grounded tool. This paradigm offers a compelling strategy to optimize perinatal management by balancing the imperative for timely intervention against the risks of iatrogenic harm [30] in this vulnerable population.

Availability of Data and Materials

All data associated with this study and the custom code developed for the AI model are available upon reasonable request from the corresponding author.

Author Contributions

JZ and YD designed the research study. JS and BL performed the research, including patient recruitment and ultrasound examinations. BL developed the AI model and performed feature extraction. JZ and JS analyzed the data and interpreted the results. YD supervised the project and provided critical revision of the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Ethical oversight for this research was managed by the Medical Ethics Committee at Wuxi People's Hospital Affiliated to Nanjing Medical University (Approval No. 2021-1624011). Written informed consent was obtained from all participants before their participation. Blood collection and data handling followed ethical standards set by the Declaration of Helsinki.

Acknowledgment

The authors would like to thank the clinical and nursing staff of the Department of Ultrasound Medicine at Wuxi People's Hospital for their assistance in patient recruitment and data collection. We also extend our gratitude to our colleagues in the IT and technical support departments for their assistance with computational infrastructure. Finally, we thank the peer reviewers and editors for their valuable comments and suggestions which have significantly improved this manuscript.

Funding

This study was supported by the Wuxi Municipal Double Hundred Young and Middle-Aged Reserve Top Talents Program in Medical and Health Fields (Grant No. HB2023001) and the Jiangsu Provincial Health Commission Scientific Research Fund Project (Grant No. X202336).

Conflict of Interest

The authors declare no conflict of interest.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the authors used ChatGpt-3.5 in order to check spell and grammar. After using this tool, the authors reviewed and edited the content

as needed and takes full responsibility for the content of the publication.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/CEOG46350.

References

[1] Mills G, Badeghiesh A, Suarthana E, Baghlaf H, Dahan MH. Polycystic ovary syndrome as an independent risk factor for gestational diabetes and hypertensive disorders of pregnancy: a population-based study on 9.1 million pregnancies. Human Reproduction (Oxford, England). 2020; 35: 1666–1674. https://doi.org/10.1093/humrep/deaa099.

[2] Huang H, Pan J, Spielberg DR, Hanchard NA, Scott DA, Burrage LC, *et al*. A dominant negative variant of *RAB5B* disrupts maturation of surfactant protein B and surfactant protein C. Proceedings of the National Academy of Sciences of the United States of America. 2022; 119: e2105228119. https://doi.org/10.1073/pnas.2105228119.

[3] Hosni A, El-Twab SA, Abdul-Hamid M, Prinsen E, AbdElgawad H, Abdel-Moneim A, *et al*. Cinnamaldehyde mitigates placental vascular dysfunction of gestational diabetes and protects from the associated fetal hypoxia by modulating placental angiogenesis, metabolic activity and oxidative stress. Pharmacological Research. 2021; 165: 105426. https://doi.org/10.1016/j.phrs.2021.105426.

[4] Besnard AE, Wirjosoekarto SAM, Broeze KA, Opmeer BC, Mol BWJ. Lecithin/sphingomyelin ratio and lamellar body count for fetal lung maturity: a meta-analysis. European Journal of Obstetrics, Gynecology, and Reproductive Biology. 2013; 169: 177–183. https://doi.org/10.1016/j.ejogrb.2013.02.013.

[5] Madan I, Romero R, Kusanovic JP, Mittal P, Chaiworapongsa T, Dong Z, *et al*. The frequency and clinical significance of intra-amniotic infection and/or inflammation in women with placenta previa and vaginal bleeding: an unexpected observation. Journal of Perinatal Medicine. 2010; 38: 275–279. https://doi.org/10.1515/jpm.2010.001.

[6] Practice Bulletin No. 162: Prenatal Diagnostic Testing for Genetic Disorders. Obstetrics and gynecology. 2016; 127: e108–e122. https://doi.org/10.1097/AOG.0000000000001405.

[7] He P, Yu L, Wang X, Ren B, Gu P, Yang F, *et al*. A prospective study of transorbital ultrasound multimodal imaging in the detection and evaluation of increased intracranial pressure. BMC Medical Imaging. 2025; 25: 349. https://doi.org/10.1186/s12880-025-01893-w.

[8] Perri A, Sbordone A, Patti ML, Nobile S, Tirone C, Giordano L, *et al*. The future of neonatal lung ultrasound: Validation of an artificial intelligence model for interpreting lung scans. A multicentre prospective diagnostic study. Pediatric Pulmonology. 2023; 58: 2610–2618. https://doi.org/10.1002/ppul.26563.

[9] Vukicevic AM, Milic V, Zabotti A, Hocevar A, De Lucia O, Filippou G, *et al*. Radiomics-Based Assessment of Primary Sjögren's Syndrome From Salivary Gland Ultrasonography Images. IEEE Journal of Biomedical and Health Informatics. 2020; 24: 835–843. https://doi.org/10.1109/JBHI.2019.2923773.

[10] Chatzakis C, Lausegger S, Sembrera E, Vargas S, Nicolaides KH, Charakida M. Maternal vascular dysfunction in gestational diabetes is associated with birth of small neonates. Diabetes Research and Clinical Practice. 2025; 221: 112032. https://doi.org/10.1016/j.diabres.2025.112032.

[11] Youssef L, Crispi F, Paolucci S, Miranda J, Lobmaier S, Crovetto F, *et al*. Angiogenic factors alone or in combination with ultrasound Doppler criteria for risk classification among late-onset small fetuses with or without pre-eclampsia. Ultrasound in Obstetrics & Gynecology. 2025; 65: 317–324. https://doi.org/10.1002/uog.29181.

[12] Gopal SH, Donepudi R, Pammi M. Leptin deficiency, a potential mechanism for impaired fetal lung development in uteroplacental insufficiency? Pediatric Research. 2024; 95: 1410–1411. https://doi.org/10.1038/s41390-024-03038-1.

[13] Meng X, Wang L, Wu M, Zhang N, Li X, Wu Q. Construction of a Prediction Model for Adverse Perinatal Outcomes in Foetal Growth Restriction Based on a Machine Learning Algorithm: A Retrospective Study. BJOG: an International Journal of Obstetrics and Gynaecology. 2025; 132: 1514–1523. https://doi.org/10.1111/1471-0528.18226.

[14] Chen W, Zeng B, Ling X, Chen C, Lai J, Lin J, *et al*. Assessing fetal lung maturity: Integration of ultrasound radiomics and deep learning. African Journal of Reproductive Health. 2025; 29: 51–64. https://doi.org/10.29063/ajrh2025/v29i5s.7.

[15] Fischhoff B. Decision science in perinatal decision making. Seminars in Perinatology. 2025; 49: 152064. https://doi.org/10.1016/j.semperi.2025.152064.

[16] Obstetrics Subgroup, Chinese Society of Obstetrics and Gynecology. Guideline of diagnosis and treatment of hyperglycemia in pregnancy (2022) [Part one]. Chinese Journal of Obstetrics and Gynecology. 2022; 57: 3–12. https://doi.org/10.3760/cma.j.cn112141-20210917-00528. (In Chinese)

[17] American Diabetes Association Professional Practice Committee. 15. Management of Diabetes in Pregnancy: Standards of Medical Care in Diabetes-2022. Diabetes Care. 2022; 45: S232–S243. https://doi.org/10.2337/dc22-S015.

[18] Xi J, Chen J, Wang Z, Ta D, Lu B, Deng X, *et al*. Simultaneous Segmentation of Fetal Hearts and Lungs for Medical Ultrasound Images via an Efficient Multi-scale Model Integrated With Attention Mechanism. Ultrasonic Imaging. 2021; 43: 308–319. https://doi.org/10.1177/01617346211042526.

[19] Demircioğlu A. Benchmarking Feature Selection Methods in Radiomics. Investigative Radiology. 2022; 57: 433–443. https://doi.org/10.1097/RLI.0000000000000855.

[20] Aguilera J, Semmler J, Coronel C, Georgiopoulos G, Simpson J, Nicolaides KH, *et al*. Paired maternal and fetal cardiac functional measurements in women with gestational diabetes mellitus at 35-36 weeks' gestation. American Journal of Obstetrics and Gynecology. 2020; 223: 574.e1–574.e15. https://doi.org/10.1016/j.ajog.2020.04.019.

[21] Xin KZ, Li D, Yi PH. Limited generalizability of deep learning algorithm for pediatric pneumonia classification on external data. Emergency Radiology. 2022; 29: 107–113. https://doi.org/10.1007/s10140-021-01954-x.

[22] Benhalima K, Yamamoto JM. Use of continuous glucose monitoring and hybrid closed-loop therapy in pregnancy. Diabetes, Obesity & Metabolism. 2024; 26 Suppl 7: 74–91. https://doi.org/10.1111/dom.15999.

[23] Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. The Spine Journal: Official Journal of the North American Spine Society. 2021; 21: 1643–1648. https://doi.org/10.1016/j.spinee.2021.02.024.

[24] Packer CH, Zhou CG, Hersh AR, Allen AJ, Hermesch AC, Caughey AB. Antenatal Corticosteroids for Pregnant Women at High Risk of Preterm Delivery with COVID-19 Infection: A Decision Analysis. American Journal of Perinatology. 2020; 37: 1015–1021. https://doi.org/10.1055/s-0040-1713145.

[25] Barrett AN, Huang Z, Aung S, Ho SSY, Roslan NS, Mahyuddin AP, *et al*. Whole-Chromosome Karyotyping of Fetal Nucleated Red Blood Cells Using the Ion Proton Sequencing Platform. Genes. 2022; 13: 2257. https://doi.org/10.3390/genes13122257.

[26] Ho JK, Safari A, Adibi A, Sin DD, Johnson K, Sadatsafavi M,

et al. Generalizability of Risk Stratification Algorithms for Exacerbations in COPD. Chest. 2023; 163: 790–798. https://doi.org/10.1016/j.chest.2022.11.041.

[27] Jelicic J, Stauffer Larsen T, Bukumiric Z, Juul-Jensen K, Andjelic B. Prognostic models in primary central nervous system lymphoma patients: A systematic review. Critical Reviews in Oncology/hematology. 2021; 161: 103341. https://doi.org/10.1016/j.critrevonc.2021.103341.

[28] Vats H, Saxena R, Sachdeva MP, Walia GK, Gupta V. Impact of maternal pre-pregnancy body mass index on maternal, fetal and neonatal adverse outcomes in the worldwide populations: A systematic review and meta-analysis. Obesity Research & Clinical Practice. 2021; 15: 536–545. https://doi.org/10.1016/j.orcp.2021.10.005.

[29] Asgari S, Khalili D, Azizi F, Hadaegh F. External validation of the American prediction model for incident type 2 diabetes in the Iranian population. BMC Medical Research Methodology. 2023; 23: 77. https://doi.org/10.1186/s12874-023-01891-y.

[30] Haizler-Cohen L, Alzamora MC, Legro NR, Eter L, Freed T, Rahematpura S, *et al*. Late Preterm Antenatal Corticosteroids in Pregestational and Gestational Diabetic Pregnancies. American Journal of Perinatology. 2025; 42: 1798–1808. https://doi.org/10.1055/a-2624-8405.