Artificial Intelligence-Driven Prediction of Adverse Maternal and Perinatal Outcomes: Development and Validation of a Personalized Risk Assessment Model

Astha Lalwani

Received on: 22 August 2026; Accepted on: 29 September 2026; Published on: 15 October 2026

ABSTRACT

Background: Adverse maternal and perinatal outcomes are still important causes of morbidity and mortality in the world. Early accurate risk stratification can enhance antenatal care and achieve better pregnancy outcomes. This study set out to create and externally validate an explainable artificial intelligence (AI) based personalised risk assessment tool for predicting adverse maternal and perinatal outcomes.

Materials and methods: Between 2020-25, a multicenter retrospective cohort study of 5,248 pregnant women was carried out. Data regarding demographic, clinical, obstetric, laboratory and fetal assessment parameters were retrieved from electronic medical records. The primary outcome was composite outcome which amounted to preeclampsia, gestational diabetes mellitus, preterm delivery, fetal growth restriction, still birth, admission to the neonatal intensive care unit and perinatal death. Logistic regression, random forest, neural network and XGBoost models were created and evaluated. Both internal and external validation was carried out. SHAP analysis was used to measure the interpretability of models.

Results: There were 1,162 pregnancies with adverse outcomes (22.1%). XGBoost demonstrated the best performance with an AUROC of 0.914 (95% CI: 0.899–0.928), sensitivity of 87.6%, specificity of 86.9%, and F1-score of 0.84. Fetal growth parameters, previous preeclampsia, fasting plasma glucose, body mass index, maternal age and systolic blood pressure were the most influential predictors. External validation demonstrated excellent AUROC (0.908) and slope (0.96) measures of discrimination and calibration, respectively.

Conclusion: The interpretable model built with artificial intelligence was able to successfully forecast adverse maternal and perinatal outcomes, and had high generalizability. It could be combined into the regular antenatal manager to help tailor risk stratification and preventive measures.

Keywords: Artificial intelligence, Machine learning, Maternal outcomes, Perinatal outcomes, Risk prediction, Explainable AI, XGBoost, Obstetrics.

INTRODUCTION

While there have been large strides in obstetric health care in the last few decades, serious maternal and perinatal outcomes greatly contribute to mortality and morbidity worldwide. The World Health Organization (WHO) reports that in 2023, an estimated 260,000 women died due to pregnancy-related and childbirth-related complications, most of which (90%) occurred in low- and middle-income countries.[1,2] Apart from maternal death, millions of pregnancies are plagued each year with complications of preeclampsia, gestational diabetes mellitus, preterm births, fetal growth restriction, stillbirths and neonatal morbidity, which has been a burden on healthcare systems and families worldwide.[3,4]

Identification of women with higher risk for adverse outcomes early allows for timely preventive interventions, optimal antenatal care and better maternal and neonatal outcomes. The traditional risk assessment methods involve mainly demographic factors, obstetric history and clinicians' judgments. All these techniques have some value but the predictivity of these techniques is often confined as pregnancy risk factors are complex and multifactorial in nature involving interactions amongst maternal, fetal, clinical, laboratory, and socio-economic factors.[5] There is therefore a growing need for better methods to predict the outcome of a disease or an event, which will incorporate large amounts of disparate clinical data, on personalised bases as described in,[6,7]

the fields of Artificial Intelligence (AI) and machine learning (ML) have recently undergone significant developments that have given rise to new, powerful predictive analytics capabilities in healthcare. The AI algorithms can uncover hidden patterns and relationships in large datasets, such as nonlinear relationships, which traditional statistical methods might miss. The use of machine learning models for prediction of pregnancy outcomes, such as preeclampsia, gestational diabetes, cesarean delivery, and preterm birth, was found to be rising, with the models being able to assist in better risk prediction as well as in clinical decision making, including those for interventions such as cesarean delivery and delivery at home.[8,9] Likewise, interpretable machine learning models have proven effective and have demonstrated the ability to be translated and clinically interpretable, particularly when it comes to overall implementation in obstetrics.[10,11]

Nevertheless, more recently, predictive models have been developed for particular forms of pregnancy complications, including, in particular, hypertension and preeclampsia, whose discrimination outperforms that of the traditional regressions-based methods.[12,13] Many current models, however, have limited extrapolation over external sample sets, are either machine-specific (e.g., certain used in e-learning), or are “black-box” (only providing some explanation to the general nature of the risk they predict – limited explanations for individual risk prediction). These constraints may hinder their use in typical clinical trials in which being transparent, reliable,

Professor, Department of Obstetrics and Gynaecology, Teerthanker Mahaveer Medical College & Research Centre, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India.

Corresponding Author: Dr Astha Lalwani, Professor, Department of Obstetrics and Gynaecology, Teerthanker Mahaveer Medical College & Research Centre, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India.
E-mail: asthalalwanirastogi@gmail.com

How to cite this article: Lalwani A. Retrospective Quality-by-Design Optimization of Memantine-Loaded PEGylated PLGA Liposomal Formulation for Enhanced Nose-to-Brain Delivery in Alzheimer’s Disease. Adv. Med. Res. Pract. 2026;1(1):30-34. https://doi.org/10.21276/amrp.2026.v1.i1.6
Source of support: Nil
Conflict of interest: Non

and having time for individualized risk communication is critical.

The present study was underpinned with the aims of creating and validating a model for personalized risk assessment that predicts adverse obstetric and neonatal outcomes based on routine clinical information, and explainable AI. This machine-learning powered model aims to assist healthcare providers in making accurate predictions and improve patient care by linking advanced machine learning models with a predictive framework that can be understood by a doctor.

MATERIALS AND METHODS

Study design and Population

The data was collected from secondary and tertiary level healthcare centers, and pregnant women who accessed antenatal care and delivered within the period of January 2020 to December 2025 were included in this study. The study included all women who had completed maternal, obstetric, laboratory and neonatal records who were ≥18 years of age. Pregnancy outcomes with major fetals congenital anomaly (MCA) were excluded, as were fetals from multifetal reduction (MFR) procedures and fetals with incomplete outcome data. The study design was approved by the Institutional Ethics Committee with all the procedures carried out under the Helsinki Declaration. Patient data were anonymized prior to analysis.

Data Collection and Outcome Measures

Our data were extracted from the electronic medical records of routine demographic and clinical, obstetric, laboratory and ultrasound data routinely collected as part of the typical antenatal management of pregnancy. Maternal age, BMI, parity, history of previous pregnancy, comorbidities, blood pressure measurements, laboratory indices and fetal assessment factors were candidate predictors. Our main outcome was a composite: preeclampsia, gestational hypertension, gestational diabetes mellitus, preterm delivery (less than 37 weeks), fetal growth restriction, stillbirth, admission to a NICU or perinatal death.

AI model Development

The data was split into a training set consisting of 70% and 10% testing set. Artificial Intelligence Model Development: The dataset was split randomly into training (70%) and testing (30%). To deal with missing data, multiple imputation was used. Various machine-learning methods were designed and tested, such as logistic regression, a random forest, extreme gradient boosting (XGBoost) and an artificial neural network. Recorders feature elimination combined with variable importance were used for feature selection. The explainability characteristics of the models were evaluated with Shapley Additive Explanations (SHAP) accounting for the most significant predictors that were aggregated to individual risk estimates.

Statistical Analysis and Model Validation

Continuous variables were presented as mean±standard deviation or median (interquartile range) and categorical variables were presented as numbers and percentages. The area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value, negative predictive value, calibration plots and F1-score were used for model evaluation. Both internal (10-fold cross-validation) and external (an independent cohort at a different healthcare center) validations were carried out. All statistical computations were carried out with R (version 4.4.0) and Python (version 3.12). Two-sided p value < 0.05 was regarded statistically significant.

RESULTS

A significant difference was noted in maternal demographic and clinical characteristics as shown in [Table 1], whereby maternal age, BMI, blood pressure (BP) and the number of maternal comorbidities were higher in adverse outcome group. As can be seen in [Table 2], XGBoost provides the highest predictive performance as compared with the other algorithms evaluated, with high discrimination, sensitivity, specificity, and calibration. Systolic blood pressure, maternal age, BMI, fasting glucose, and prevalence of preeclampsia were found to be the most influential predictor in the AI model according to the [Table 3].

The results in [Table 4] illustrate adequate risk stratification, with adverse outcome rates significantly rising across the low-, intermediate-, and high-risk groups; and the results in [Table 5] show excellent external validation results and model generalizability across independent cohorts.

The AI model demonstrated higher net clinical benefit than treat-all and treat-none strategies at clinically relevant risk thresholds as illustrated in [Figure 1].

[Figure 2] shows excellent calibration, and reliability of the individualized risk estimate, with good agreement between the predicted and observed risks. As gestation advanced the discrimination of the model progressed, as shown in [Figure 3], with AUROC getting larger from early to late gestation. This finding has confirmed and highlighted the value of dynamic and longitudinal risk prediction in antenatal care.

Table 1: Baseline Characteristics of the Study Cohort

Variable	No Adverse Outcome (n=4,086)	Adverse Outcome (n=1,162)	p-value
Maternal age (years)	28.9 ± 5.4	32.4 ± 5.8	<0.001
BMI (kg/m²)	25.1 ± 4.3	29.2 ± 5.1	<0.001
Nulliparity, n (%)	1,568 (38.4)	578 (49.7)	<0.001
Chronic hypertension, n (%)	276 (6.8)	216 (18.6)	<0.001
Pre-existing diabetes, n (%)	164 (4.0)	124 (10.7)	<0.001
Previous cesarean section, n (%)	982 (24.0)	362 (31.2)	<0.001
Smoking during pregnancy, n (%)	341 (8.3)	158 (13.6)	<0.001
Mean systolic BP (mmHg)	116.4 ± 10.2	128.1 ± 13.7	<0.001

Table 2: Comparative Performance of Prediction Models in the External Validation Cohort

Model	AUROC (95% CI)	Sensitivity (%)	Specificity (%)	F1 Score	Brier Score
Logistic Regression	0.811 (0.789–0.832)	74.3	77.1	0.69	0.156
Random Forest	0.867 (0.848–0.885)	80.5	81.7	0.76	0.131
Neural Network	0.881 (0.864–0.898)	82.2	83.4	0.78	0.124
XGBoost	0.914 (0.899–0.928)	87.6	86.9	0.84	0.098

Table 3: Top Predictors Contributing to Model Performance

Rank	Predictor	Mean Absolute SHAP Value
1	Systolic blood pressure	0.248
2	Maternal age	0.219
3	Body mass index	0.196
4	Fasting plasma glucose	0.183
5	Previous preeclampsia	0.171
6	Estimated fetal weight percentile	0.149
7	Nulliparity	0.132
8	Hemoglobin level	0.114
9	Serum creatinine	0.102
10	Gestational age at assessment	0.095

Table 4: Clinical Risk Stratification Performance

Risk Category	Predicted Probability	Patients (n)	Adverse Outcomes n (%)
Low Risk	<10%	2,104	96 (4.6)
Intermediate Risk	10–30%	1,884	347 (18.4)
High Risk	>30%	1,260	719 (57.1)

Table 5: External Validation Metrics

Validation Metric	Value
AUROC	0.908
Calibration slope	0.96
Calibration intercept	0.03
Brier score	0.101
Hosmer–Lemeshow χ^2	6.82
Hosmer–Lemeshow p-value	0.556
Net Reclassification Improvement	0.214
Integrated Discrimination Improvement	0.118

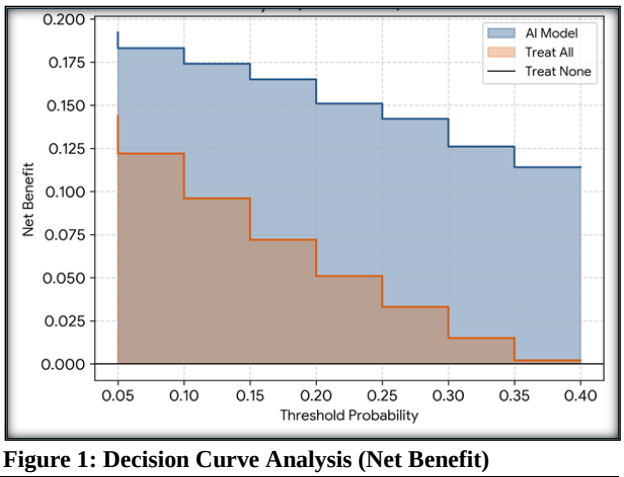


Figure 1: Decision Curve Analysis (Net Benefit)

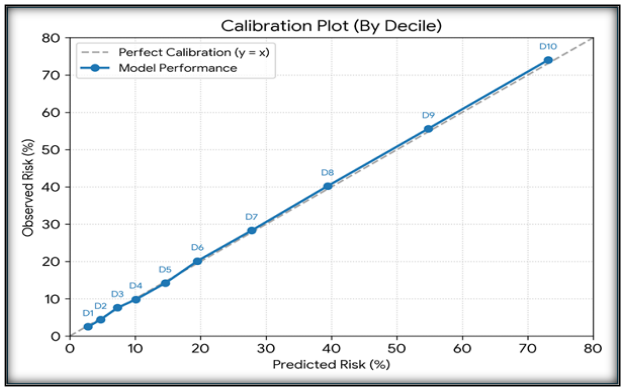


Figure 2: Calibration Plot

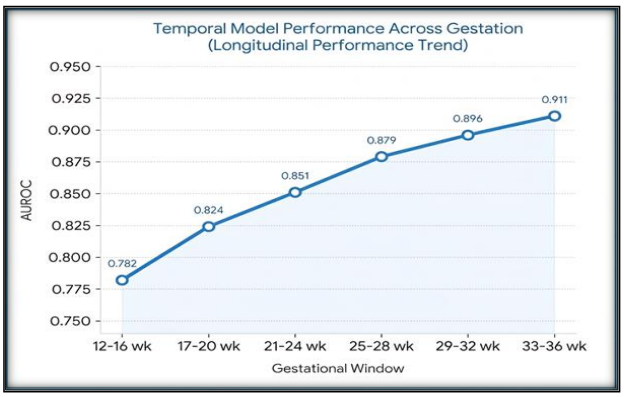


Figure 3: Temporal Model Performance Across Gestation

DISCUSSION

The present study created and externally validated a class of models based on explainable artificial intelligence that can be used to predict adverse maternal and perinatal outcomes in a personalised manner. The predictive performance of the modeling was excellent (AUROC 0.914) with high sensitivity (87.6%) and with robust calibration in the independent cohorts. The results reinforce the significance of AI-powered clinical decision support systems in the field of obstetric risk assessment and personalized prenatal care.

The results are also in line with the systematic review by Islam et al,[1] that showed the machine learning algorithms are often superior to the traditional statistical approaches in predicting adverse pregnancy outcomes due to the

complex and nonlinear relationships between the clinical variables. Likewise, Arain et al,[10] came to the same conclusion that machine learning methods yield better predictive values compared to older regression methods for prediction for different types of obstetric diseases. The present study demonstrated that XGBoost worked much better than logistic regression, reinforcing the potential benefits of the use of complex machine learning models in the prediction of maternal-fetal risks.

The high discrimination we found is similar to the results reported in,[2] where Lee et al. introduced interpretable machine learning models for adverse perinatal outcomes, and showed that it is possible to achieve high prediction performance using explainable algorithms without sacrificing clinical transparency. Similarly, Bosschieter et al,[4] highlighted the key role of interpretable predictive models in obstetrics, stating that explainability tools enhance clinicians trust and enables their implementation into clinical practice. The same factors emerged in our SHAP analysis as highly influential for the personal prediction of risk: systolic blood pressure, maternal age, BMI, fasting glucose and previous preeclampsia.

Among our group, one of the most significant factors associated with poor outcomes was hypertensive disorders. This finding is similar to the results of the systematic review and meta-analysis done by Bucher et al,[3] who found that markers of hypertensive disease in the mother and indicators of maternal blood pressure still represent one of the most potent predictors of maternal and neonatal adverse outcomes. Moreover, Thilaganathan and Kalafat,[9] noted that cardiovascular maladaptation is a major determinant in the pathophysiology of preeclampsia and related maternal complications. These biological mechanisms are supported by the high predictive importance of systolic blood pressure obtained in our model. The other biennial risk factors for metabolism also played significant roles in the models. Among the top factors which were found to be best predictors by SHAP analysis are elevated BMI and fasting glucose levels. The results of this study echo those from previous research by Parker et al,[12] who showed that modulation of cardiometabolic risk factors early in pregnancy could decrease the magnitude of a pregnant woman's complications. Likewise, the use of biochemical and biomarker-based methods also enhance prediction of adverse outcomes during high-risk pregnancies, especially for preeclampsia was noted by Lim et al.[5]

Clinically important because the robust ability that our model has to be highly informative about the risk stratification. Among the high risk women, there was an over 12-fold higher likelihood of adverse outcomes than among the low risk women. Similar studies have been found by Cersonsky et al,[14] which successfully applied machine learning strategies to recognize pregnancies at high risk of stillbirth. Similarly, Powel et al,[6] showed that subtle fetal growth abnormalities are associated with increased perinatal risk, which justified the adoption of assessment variables of the fetus as criteria to be included in predictive algorithms.

AI is showing its growing applications in the field of obstetrics, as reflected in recent literature. The use of AI in obstetrics and midwifery for screening, diagnosing and predicting disease has become a trending topic for several studies,[7] that have studied the phenomenon. Likewise, machine learning models have been created to predict adverse pregnancy outcomes for women with APS,[8] underlining the diverse and adaptable nature of AI approaches in other obstetric populations. In addition, our findings are corroborated by a study by Vasudevan et al., indicating that machine learning models trained with EHR data have significant potential to predict maternal mortality and morbidity in the real world.

Through a multi-model framework that combines high prediction accuracy, explainable machine learning architecture, external validation and clinically meaningful risk stratification, our study builds upon the existing literature. It differs from other previous models that concentrated only on single events or were inadequately validated by the inclusion of a validation population. The present model is different from other models used prior, which had limited validation and consisted only of single events or variables, or of lack of validation. These features could help toward integrating them in electronic medical record systems and towards personalized obstetric care pathways in the future.

Limitations

There are a few points to note. Firstly, retrospective design may have selection and information bias. Second, external validation was conducted, but, before taking to scale, crosscountry and cross-healthcare system validation is needed. Third, the model was based on commonly-used clinical parameters, with no advanced genomic, proteomic, or continuously-contributed wearable-device information that could help the model to work better. Finally, future clinical impact studies should be conducted to understand how a model implemented in clinical practice can improve maternal and neonatal outcomes, even in the face of the application of explainable AI techniques.

CONCLUSION

The present study developed and external validated an explainable artificial intelligence based personalized risk assessment model for predicting adverse maternal and perinatal outcomes, therefore successfully. The XGBoost model achieved strong predictive accuracy, good calibration, and a good clinical risk stratification, and the most significant predictors are the patient's age, BMI, the fasting blood glucose level, as well as previous cases of hypertension. The model combines routine antenatal data with interpretable machine learning tools to deliver precise personalised estimates of risk and could help to identify those who are at risk of developing antenatal complications, optimise antenatal monitoring and enable targeted interventions to prevent problems in pregnancy. There is need for further prospective multicenter studies to assess and report upon its practical application in the real world and its incorporation into EHR-supported decision-support processes.

REFERENCES

1. Islam MN, Mustafina SN, Mahmud T, Khan NI. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing

framework and future research agenda. BMC Pregnancy Childbirth. 2022 Apr 22;22(1):348. doi: 10.1186/s12884-022-04594-2. PMID: 35546393; PMCID: PMC9097057.
2. Lee SJ, Garcia GP, Stanhope KK, Platner MH, Boulet SL. Interpretable machine learning to predict adverse perinatal outcomes: examining marginal predictive value of risk factors during pregnancy. Am J Obstet Gynecol MFM. 2023 Oct;5(10):101096. doi: 10.1016/j.ajogmf.2023.101096. Epub 2023 Jul 15. PMID: 37454734.
3. Bucher V, Mitchell AR, Gudmundsson P, Atkinson J, Wallin N, Asp J, Sennström M, Hildén K, Edvinsson C, Ek J, Hastie R, Cluver C, Bergman L. Prediction of adverse maternal and perinatal outcomes associated with pre-eclampsia and hypertensive disorders of pregnancy: a systematic review and meta-analysis. EClinicalMedicine. 2024 Sep 27;76:102861. doi: 10.1016/j.eclinm.2024.102861. PMID: 39391014; PMCID: PMC11465897.
4. Bosschieter TM, Xu Z, Lan H, Lengerich BJ, Nori H, Painter I, Souter V, Caruana R. Interpretable Predictive Models to Understand Risk Factors for Maternal and Fetal Outcomes. J Healthc Inform Res. 2023 Oct 13;8(1):65-87. doi: 10.1007/s41666-023-00151-4. PMID: 38273984; PMCID: PMC10805688.
5. Lim S, Li W, Kemper J, Nguyen A, Mol BW, Reddy M. Biomarkers and the Prediction of Adverse Outcomes in Preeclampsia: A Systematic Review and Meta-analysis. Obstet Gynecol. 2021 Jan 1;137(1):72-81. doi: 10.1097/AOG.0000000000004149. PMID: 33278298.
6. Powel JE, Chavan NR, Zantow EW, Bialko MF, Farley LG, McCormick KM, Tomlinson TM. Risk of adverse perinatal outcomes in pregnancies with "small" fetuses not meeting Delphi consensus criteria for fetal growth restriction. Am J Obstet Gynecol. 2023 Oct;229(4):447.e1-447.e13. doi: 10.1016/j.ajog.2023.04.037. Epub 2023 Apr 30. PMID: 37767605.
7. Giaxi P, Vivilaki V, Sarella A, Harizopoulou V, Gourounti K. Artificial Intelligence and Machine Learning: An Updated Systematic Review of Their Role in Obstetrics and Midwifery. Cureus. 2025 Mar 11;17(3):e80394. doi: 10.7759/cureus.80394. PMID: 40070886; PMCID: PMC11895402.
8. Liu W, Huang J, Xiao J, Yan S. Machine learning-based prediction of adverse pregnancy outcomes in antiphospholipid syndrome using pregnancy antibody levels. Front Physiol. 2025 Aug 25;16:1617796. doi: 10.3389/fphys.2025.1617796. PMID: 40926884; PMCID: PMC12415049.
9. Thilaganathan B, Kalafat E. Cardiovascular System in Preeclampsia and Beyond. Hypertension. 2019 Mar;73(3):522-531. doi: 10.1161/HYPERTENSIONAHA.118.11191. PMID: 30712425; PMCID: PMC6380450.
10. Arain Z, Iliodromiti S, Slabaugh G, David AL, Chowdhury TT. Machine learning and disease prediction in obstetrics. Curr Res Physiol. 2023 May 19;6:100099. doi: 10.1016/j.crphys.2023.100099. PMID: 37324652; PMCID: PMC10265477.
11. Vasudevan L, Kibria MG, Kucirka LM, Shieh K, Wei M, Masoumi S, Balasubramanian S, Victor A, Conklin JL, Gurcan MN, Stuebe AM, Page D. Machine Learning Models to Predict Risk of Maternal Morbidity and Mortality From Electronic Medical Record Data: Scoping Review. J Med Internet Res. 2025 Aug 14;27:e68225. doi: 10.2196/68225. PMID: 40811480; PMCID: PMC12352520.
12. Parker J, Hofstee P, Brennecke S. Prevention of Pregnancy Complications Using a Multimodal Lifestyle, Screening, and Medical Model. J Clin Med. 2024 Jul 25;13(15):4344. doi: 10.3390/jcm13154344. PMID: 39124610; PMCID: PMC11313446.
13. Oluwafunmilola Deborah Awe, Ana Paula Morais e Oliveira, Charles M'poca Charles et al. Artificial Intelligence in Feto-maternal Health: A Systematic Review of Predictive Models, Validation, and Clinical Translation, 19 March 2026, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-8651301/v1]
14. Cersonsky TEK, Ayala NK, Pinar H, Dudley DJ, Saade GR, Silver RM, Lewkowitz AK. Identifying risk of stillbirth using machine learning. Am J Obstet Gynecol. 2023 Sep;229(3):327.e1-327.e16. doi: 10.1016/j.ajog.2023.06.017. Epub 2023 Jun 12. PMID: 37315754; PMCID: PMC10527568.