European Journal of Integrative Medicine 81 (2026) 102596

Contents lists available at ScienceDirect

European Journal of Integrative Medicine

journal homepage: www.sciencedirect.com/journal/european-journal-of-integrative-medicine

Using explainable machine learning to identify predictors of Kangaroo mother care implementation in Sierra Leone's healthcare system

Afeez A. Soladoye [a,b], David B. Olawade [c,d,e,*], Joseph E. Origbo [f], Kobloobase O. Usani [g], Ayomide Israel Adekoya [h,i], Ojima Z. Wada [j], Augustus Osborne [k]

[a] Department of Computer Engineering, Federal University, Oye-Ekiti, Nigeria
[b] Department of Computer Engineering, Adeleke University, Ede, Nigeria
[c] Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, UK
[d] Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham ME7 5NY, UK
[e] Department of Business, Management and Health, York St John University, London, E14 2BA, United Kingdom
[f] Department of Computer Science, Nottingham Trent University, Nottingham, UK
[g] Department of Data Science and Artificial Intelligence, School of Business, Computing and Social Sciences, University of Gloucestershire, Cheltenham, UK
[h] Department of Computer Science, Faculty of Computing, Engineering and the Built Environment, Birmingham City University, Birmingham, UK
[i] Department of Physiology, Faculty of Basic Medical Sciences, University of Ibadan, Ibadan, Nigeria
[j] College of Science and Engineering, Division of Sustainable Development, Hamad Bin Khalifa University, Qatar Foundation, Education City, Doha, Qatar
[k] Institute for Development, Western Area, Freetown, Sierra Leone

ARTICLE INFO

Keywords:
Kangaroo Mother Care
Machine Learning
Explainable Artificial Intelligence
Maternal Health
Sierra Leone

ABSTRACT

Introduction: Kangaroo Mother Care (KMC) reduces neonatal mortality and improves thermoregulation and breastfeeding, yet uptake remains inconsistent in Sierra Leone. Predictive and explainable tools could target implementation where the need is most significant and resources are scarce. This study aimed to predict KMC adoption and identify actionable predictors using explainable machine learning.

Methods: We analysed a nationally representative dataset from Sierra Leone comprising 7737 births. The study setting was Sierra Leone's healthcare system, with participants including mothers who delivered in health facilities. Following data preprocessing (imputation, MinMax normalisation, categorical encoding, and SMOTE for class imbalance), forward-backward selection reduced 22 candidate variables to 10 key predictors. Five classifiers were trained using a 70:30 stratified split: K-Nearest Neighbors (KNN), logistic regression (LR), Support Vector Machine (SVM), Random Forest (RF), and XGBoost. The outcome was KMC adoption (binary: received/not received). Performance was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Interpretability was achieved through SHAP and LIME for global and local explanations.

Results: XGBoost performed best (accuracy 0.72, precision 0.75, recall 0.81, F1 0.78, ROC AUC 0.7685), followed by Random Forest. Predictors associated with KMC included delivery by caesarean section, type of birth, maternal employment, number of antenatal visits, place of delivery, health insurance coverage, and region, while sampling design variables captured contextual heterogeneity. SHAP and LIME consistently highlighted delivery characteristics and socio-economic factors as primary drivers.

Conclusion: Explainable ensemble models can flag infants likely to receive or miss KMC and indicate modifiable levers for improvement. High recall supports use as a screening aid to prioritise counselling, facility preparedness, and postnatal support. Prospective validation, threshold calibration, and integration within routine health information systems are warranted to translate these insights into sustained increases in KMC coverage in Sierra Leone and similar settings.

Abbreviations: AI, Artificial Intelligence; AUC, Area Under the Curve; AUROC, Area Under the Receiver Operating Characteristic Curve; EEG, Electroencephalogram; GNN, Graph Neural Network; KMC, Kangaroo Mother Care; KNN, K-Nearest Neighbors; LIME, Local Interpretable Model-agnostic Explanations; LR, Logistic Regression; ML, Machine Learning; NICU, Neonatal Intensive Care Unit; RBF, Radial Basis Function; RF, Random Forest; ROC, Receiver Operating Characteristic; SHAP, SHapley Additive exPlanations; SMOTE, Synthetic Minority Oversampling Technique; SVM, Support Vector Machine; TSB, Total Serum Bilirubin; WHO, World Health Organization; XAI, Explainable Artificial Intelligence.

* Corresponding author.
E-mail address: d.olawade@uel.ac.uk (D.B. Olawade).

https://doi.org/10.1016/j.eujim.2025.102596
Received 24 September 2025; Received in revised form 29 November 2025; Accepted 5 December 2025
Available online 9 December 2025
1876-3820/

1. Introduction

Kangaroo Mother Care (KMC) is one of the simplest and most effective ways to save the lives of babies born small or early. It involves placing a newborn directly on the mother's bare chest, supporting breastfeeding, and encouraging early discharge when safe. Decades of research show that this low-cost method helps babies stay warm, gain weight, avoid infections, and bond with their mothers [1,2]. In 2023, the World Health Organization reaffirmed KMC as the standard of care for these vulnerable infants and recommended starting it immediately after birth [3]. Despite these substantial benefits, many babies who could receive KMC still do not, especially in countries with limited healthcare resources [4,5]. This gap is especially concerning in Sub-Saharan Africa [6], where newborn deaths remain among the highest in the world [7,8]. Sierra Leone, with its unique healthcare landscape shaped by recent health system strengthening efforts following the Ebola epidemic, presents a critical case study for examining KMC implementation patterns [9].

In recent years, researchers and healthcare workers have increasingly explored how artificial intelligence (AI) and machine learning (ML), computer systems that learn patterns from data, can support maternal and newborn health [10–14]. These tools can help identify mothers and infants at risk and enable health systems to act earlier. For example, advanced computer models have been used to detect neonatal jaundice from images [29], identify seizures using brainwave recordings [30], and recognize newborns to prevent accidental swapping in hospitals [31,32]. ML can also predict important events during pregnancy, such as emergency caesarean sections [33], restricted fetal growth [34], non-reassuring fetal heart patterns [35], the likelihood of episiotomy [36], and whether early skin-to-skin contact will happen after birth [37]. These examples show how AI can support health workers, especially in places where resources and staff are limited.

AI tools are most helpful when their decisions are easy to understand. Explainable Artificial Intelligence (XAI) helps achieve this by showing how and why a model makes specific predictions [15]. Two widely used XAI methods are SHAP and LIME, which highlight the factors influencing each prediction in transparent and interpretable ways [16–20]. This transparency is essential in healthcare, where trust, clarity, and accountability matter. Combining accurate predictions with explainable results can help health workers see not just who is at risk, but also which factors contribute most to those risks.

However, little is known about how these techniques can help improve KMC adoption in real health systems. Many factors: clinical, economic, social, and regional, may influence whether a newborn receives KMC [21–23]. Understanding these factors through a data-driven and explainable approach can support better planning, guide counselling for mothers, strengthen health facilities, and reduce preventable newborn deaths.

In this study, we aimed to use ML to predict which newborns in Sierra Leone receive Kangaroo Mother Care and to identify the most important factors influencing this decision. We analysed a nationally representative dataset of 7737 births using multiple machine learning models. We applied explainable AI methods (SHAP and LIME) to ensure that the findings are clear, transparent, and valid for healthcare workers and policymakers. Our goal is to provide evidence that can support practical actions to increase KMC adoption and improve newborn survival in resource-limited settings.

2. Methods

2.1. Data acquisition and description

This study employed a comprehensive preprocessing pipeline to ensure the dataset was suitable for training machine learning models to predict KMC adoption in children born in Sierra Leone. A series of preprocessing steps was applied to handle missing values, normalise numerical features, encode categorical variables, and address class imbalance. These steps were critical to preparing a robust dataset that would support accurate, reliable model predictions. The dataset used in this study is a nationally representative sample of births in Sierra Leone, comprising 7737 births. Data were collected from mothers who delivered in health facilities across the country, capturing a wide range of clinical, socio-economic, and demographic variables. This large sample size provides sufficient statistical power for machine learning model development and evaluation, enabling robust predictions of KMC adoption.

2.2. Data preprocessing

I **Handling Missing Values:** Missing values in the dataset were imputed using the mean for each numerical column. This approach was chosen because it preserves the data's central tendency and minimises bias, particularly when missingness is assumed to be random. For each feature with missing values, the mean was calculated using available data points within the same column, and missing entries were replaced with this value. This method ensured the dataset remained complete without discarding valuable records or imposing complex assumptions about missing-data patterns.

II **Normalisation:** To ensure that numerical features with varying scales did not unduly influence model performance, MinMax normalisation was applied. This technique rescales feature values to a fixed range of [0, 1] using the formula:

$$X_{Norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where X is the original feature value, X_min is the minimum value in the feature column, and X_max is the maximum value. MinMax normalisation was selected because it preserves the relative relationships between data points whilst ensuring compatibility with algorithms sensitive to feature scales, such as Support Vector Machines (SVMs) and K-Nearest Neighbours (KNNs).

In the MinMax normalisation formula, (X) represents the original feature value for a given data point, (X_{min}) is the minimum value observed in that feature column across the dataset, and (X_{max}) is the maximum value in the same column. This transformation rescales the feature values to the fixed range [0, 1], preserving the relative relationships between data points while ensuring compatibility with machine learning algorithms sensitive to feature scales, such as Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs).

III. **Encoding Categorical Variables:** Categorical variables in the dataset were encoded as numerical values to make them compatible with machine learning algorithms. The .replace function was used to map categorical values to numerical equivalents based on predefined mappings. For instance, ordinal categories (e.g., Poorer, Poor, Middle, Richer and Richest) were assigned sequential integers (1, 2, 3, 4 and 5), whilst nominal categories (such as Region, Frequency of listening to radio, Type of place of residence and others) were encoded using distinct integers. This approach was chosen over one-hot encoding to reduce dimensionality and over label encoding to allow the researcher to assign distinct integers to corresponding variables, as the dataset contained a moderate number of categorical variables, and the selected models (e.g., XGBoost, Random Forest) are robust to non-binary numerical encodings.

IV. **Addressing Class Imbalance:** The dataset was imbalanced, with a disproportionate number of instances in the non-KMC class (2950) compared to the KMC class (4787). To mitigate this, the Synthetic Minority Oversampling Technique (SMOTE) was applied. SMOTE generates synthetic samples for the minority class by interpolating between existing minority class instances, thereby balancing the class distribution without introducing excessive noise. This technique was selected because it improves model performance on imbalanced datasets by reducing bias toward the majority class, which is particularly important for health-related predictions where correctly identifying the minority class is critical.

These preprocessing steps collectively ensured that the dataset was clean, normalised, and balanced, providing a solid foundation for training and evaluating machine learning models.

2.3. Feature selection

To enhance model performance and reduce computational complexity, feature selection was performed to identify the most relevant predictors of KMC adoption. The dataset initially contained 22 features, encompassing demographic, clinical, and socio-economic variables. Forward-backward feature selection, a hybrid wrapper method, was employed to select an optimal subset of 10 features. Forward-backward feature selection combines forward selection (iteratively adding features that improve model performance) and backward elimination (removing features that contribute least to performance). This method was chosen because it balances feature-combination exploration with computational efficiency, making it suitable for datasets with a moderate number of features. The selection process used a performance metric (e.g., accuracy or F1 score) on a validation set to evaluate feature subsets, iteratively adding or removing features until the optimal subset was identified.

The final 10 features selected through this process include delivery by caesarean section, type of birth, maternal employment status (respondent currently working), number of antenatal visits during pregnancy, place of delivery, health insurance coverage, region, women's individual sample weight, primary sampling unit, and sample strata for sample errors. These features were chosen based on their significant contribution to model performance, as evaluated by metrics such as accuracy or F1 score on a validation set. By reducing the feature set from 22 to 10, this approach minimised overfitting, improved model interpretability, and reduced training time, whilst retaining the most informative predictors.

2.4. Prediction of KMC using machine learning models

To predict KMC adoption, five ML models were employed: Random Forest (RF), XGBoost, SVM, Logistic Regression (LR), and KNN. These models were chosen for their complementary strengths in handling diverse data characteristics, robustness across different feature types, and suitability for binary classification tasks in healthcare settings.

I **Random Forest:** Random Forest is an ensemble learning method selected for its ability to handle complex, non-linear relationships in data and its robustness to overfitting. By constructing multiple decision trees and aggregating their predictions, Random Forest provides stable, accurate predictions, making it well-suited to datasets with mixed feature types, such as those in this study.

II **XGBoost:** XGBoost is a gradient boosting algorithm and was chosen as the primary model due to its superior performance on structured data tasks and its effective handling of imbalanced datasets after SMOTE preprocessing. XGBoost optimises a loss function through iterative boosting, incorporating regularisation to prevent overfitting, which is particularly valuable for small-to-medium-sized healthcare datasets like the one used in this study.

III **Support Vector Machine (SVM):** SVM was included to explore its capability to find an optimal hyperplane for separating classes in high-dimensional spaces. With a radial basis function (RBF) kernel, SVM can capture non-linear patterns, making it a strong candidate for this task. Its sensitivity to feature scaling was addressed through MinMax normalisation, ensuring optimal performance.

IV **Logistic Regression (LR):** Logistic Regression was also selected as a baseline model due to its simplicity, interpretability, and effectiveness in binary classification tasks. Despite its linearity assumption, LR provides a robust benchmark for comparing the performance of more complex models and is widely used in medical research for its straightforward interpretation of coefficients.

V **K-Nearest Neighbors (KNN):** KNN was chosen for its non-parametric nature, which allows it to capture local patterns in the data without assuming a specific functional form. Its performance depends on proper feature scaling, which was ensured through MinMax normalisation. KNN was included to assess whether local neighborhood-based predictions could complement the global patterns captured by other models.

These models were selected to provide a comprehensive evaluation of KMC prediction, balancing interpretability, robustness, and predictive power. Their diverse approaches to classification ensured the study captured a wide range of patterns in the data, increasing confidence in the results' reliability.

2.5. Hyperparameter optimisation

To enhance the performance of the selected ML models, a systematic hyperparameter optimisation process was conducted using grid search, a widely accepted method for tuning ML models. Grid search exhaustively evaluates a predefined set of hyperparameter combinations to identify the configuration that maximises model performance, ensuring a robust and reproducible optimisation process. The implementation used the GridSearchCV function from the scikit-learn library in Python.

The hyperparameter grid for XGBoost included n_estimators (number of trees) with values of 100 and 200, max_depth (maximum depth of each tree) with values of 3 and 5, learning_rate (step size shrinkage) with values of 0.01 and 0.1, and subsample (fraction of samples used per tree) with values of 0.8 and 1.0. This grid was selected to balance computational efficiency with the need to explore a range of configurations that could influence model complexity and generalisation. The optimisation process used a 3-fold cross-validation strategy with the weighted F1 score as the evaluation metric, prioritising balanced performance given the dataset's initial class imbalance, which SMOTE addressed. The grid search identified the following optimal hyperparameters for the XGBoost model: learning_rate=0.1, max_depth=5, n_estimators=200, and subsample=1.0. These parameters suggest a model with a moderate learning rate to ensure stable convergence, a maximum depth of 5 to control tree complexity and prevent overfitting, 200 estimators to capture sufficient patterns in the data, and full subsampling to utilise all available data per tree. The optimised model was subsequently used for training and evaluation, ensuring that the XGBoost predictions were based on a configuration tailored to the dataset's characteristics.

2.6. Explainability methodology

To enhance the interpretability of the XGBoost model, which was identified as the primary model due to its superior predictive performance, two state-of-the-art explainability techniques were employed: SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). These methods were used to provide both global and instance-level insights into the model's predictions, ensuring transparency and generalisability, which are critical for healthcare applications where trust and understanding are paramount.

I. **SHAP:** SHAP, based on cooperative game theory, was used to quantify each feature's contribution to the model's predictions [17]. The

SHAP summary plot provides a global view of feature importance, showing the average impact of each feature on the model's output across all instances in the dataset. This plot ranked the 10 selected features by their mean absolute SHAP values, highlighting which factors (e.g., maternal education, infant health status) most strongly influenced KMC adoption predictions. The summary plot also visualised the directionality of feature effects, showing whether higher or lower feature values were associated with increased likelihood of KMC adoption.

For instance-level interpretation, SHAP waterfall plots were generated for multiple randomly selected instances from the test set. These plots decomposed individual predictions into contributions from each feature, showing how the model arrived at a specific output for a given child. By examining several waterfall plots, patterns in feature contributions were identified, enabling an assessment of the model's consistency and generalisability across diverse cases. This approach ensured transparency by providing clear, visual explanations of how specific features drove predictions for individual instances, which is particularly valuable for healthcare practitioners seeking to understand model decisions in Sierra Leone's healthcare system.

II. **LIME:** To complement SHAP's explanations, LIME was employed to provide local, model-agnostic interpretations [16]. LIME approximates the complex XGBoost model with a simpler, interpretable model (e.g., linear regression) near a specific instance, highlighting the features most influential for that prediction. LIME's output was visualised as bar plots, showing the relative importance of features for selected instances. These plots enriched the SHAP analysis by offering an alternative perspective on feature contributions, reinforcing the robustness of the interpretability framework. By comparing LIME's local explanations with SHAP's global and instance-level insights, the study ensured a comprehensive understanding of the model's behavior, addressing potential limitations of any single explainability method [20].

The combination of SHAP and LIME was chosen because it balances global interpretability (via SHAP summary plots) with local, instance-specific insights (via SHAP waterfall plots and LIME bar plots). This dual approach enhances transparency, facilitates validation of the model's decision-making process, and builds trust among stakeholders, such as healthcare providers and policymakers, by making the XGBoost model's predictions more interpretable and actionable in the context of KMC adoption.

2.7. Evaluation method and metrics

To assess the performance of the machine learning models, a hold-out evaluation method was used, with the dataset split into 70 % for training and 30 % for testing. This approach was chosen over k-fold cross-validation due to the relatively small dataset size and the need to maintain a sufficiently large test set to evaluate model generalisation in a real-world setting. Stratified sampling was applied during the split to keep the same class distribution in both sets, addressing the class imbalance handled by SMOTE during training. The 70–30 split ensured that the training set was large enough to capture the underlying patterns in the data, whilst the test set was sufficiently robust to provide reliable performance estimates. All performance metrics, including accuracy, precision, recall, F1 score, and ROC-AUC, were computed on the hold-out test set to ensure that the reported results reflect the models' ability to generalize to unseen data.

Model performance was evaluated using four metrics: accuracy, precision, recall, and F1 score. These metrics were selected to provide a comprehensive assessment of model performance, particularly in the context of imbalanced healthcare data:

- **Accuracy:** Measures the proportion of correct predictions and provides an overall assessment of model performance.
- **Precision:** Quantifies the proportion of positive (KMC) predictions that were correct, which is critical for minimising false positives in healthcare settings where incorrect KMC predictions could lead to misallocated resources.
- **Recall:** Measures the proportion of actual KMC cases correctly identified, ensuring that the model captures as many true KMC cases as possible, which is vital for maximising health outcomes.
- **F1 Score:** The harmonic mean of precision and recall, providing a balanced measure of model performance when class imbalance is a concern.

These metrics collectively ensured a thorough evaluation of the models' ability to accurately and reliably predict KMC adoption, aligning with the needs of healthcare practitioners and policymakers in Sierra Leone.

3. Results

This study presents the results of applying machine learning techniques to predict KMC adoption among children born in Sierra Leone, leveraging a dataset enriched through meticulous preprocessing and feature selection. The results highlight the performance of multiple models, the key features driving predictions, and the implications for identifying children likely to receive KMC.

3.1. Identified features from forward-backward feature selection

The forward-backward feature selection technique identified 10 out of the initial 22 features as the most influential predictors of KMC adoption. These features, selected based on their contribution to model performance, include: Delivery by Caesarean section, Type of birth, Respondent currently working, Number of antenatal visits during pregnancy, Place of delivery, Covered by health insurance, Region, Women's individual sample weight (6 decimals), Primary sampling unit, and Sample strata for sample errors. This subset reflects a combination of clinical, socio-economic, and demographic factors that collectively shape KMC uptake.

The inclusion of Delivery by Caesarean section and Type of birth underscores the role of obstetric practices in influencing KMC adoption, potentially due to variations in postpartum care protocols. Socio-economic indicators such as Respondent currently working and covered by health insurance highlight the financial and logistical barriers or enablers to KMC, suggesting that maternal employment status and insurance coverage may facilitate access to healthcare resources supportive of KMC. The Number of antenatal visits during pregnancy and the Place of delivery emphasise the importance of prenatal care and institutional delivery settings, which are critical for KMC implementation. Region, along with sampling-related variables (Women's individual sample weight, Primary sampling unit, and Sample strata for sample errors), accounts for geographical and statistical heterogeneity, ensuring the model captures regional disparities and sampling biases inherent in the dataset. These findings align with prior research on maternal and child health, reinforcing the multidimensional nature of KMC adoption and providing a robust feature set for subsequent modelling (Smith et al., 2017).

3.2. Performance of machine learning models

The predictive performance of the five machine learning models: KNN, LR, SVM with RBF kernel, RF, and XGBoost, was evaluated using a 70–30 hold-out split, with metrics including accuracy, precision, recall, F1-score, and ROC-AUC. The results are summarised in Table 1.

KNN achieved an accuracy of 0.61, a precision of 0.70, a recall of 0.60, an F1-score of 0.65, and an ROC-AUC of 0.66, indicating moderate performance and suggesting limitations in capturing the dataset's complex patterns. This outcome likely stems from KNN's reliance on local data structures, which may not generalise effectively given the heterogeneity introduced by regional and socio-economic factors, setting a baseline for comparison with more advanced models.

Transitioning to LR, the model achieved 0.68 in accuracy, 0.68 in precision, 0.88 in recall, 0.77 in F1-score, and 0.6285 in ROC-AUC, demonstrating a strong recall and sensitivity in identifying KMC cases. However, this high recall comes with a trade-off: lower precision and ROC-AUC suggest an increase in false positives, a limitation attributable to the model's linear assumptions that may not fully capture the non-linear dynamics in the data. This performance bridges to the SVM with RBF kernel, which yielded an accuracy of 0.67, precision of 0.71, recall of 0.76, F1-score of 0.73, and ROC-AUC of 0.6949, reflecting a balanced capability to model non-linear relationships. Despite this improvement over LR, SVM's overall efficacy remains below that of ensemble methods, indicating that whilst it addresses some complexity, it lacks the comprehensive robustness needed for optimal prediction.

The superior performance of ensemble techniques becomes evident with Random Forest, which achieved an accuracy of 0.72, precision of 0.75, recall of 0.80, F1-score of 0.78, and ROC-AUC of 0.7689, underscoring the strength of its approach in capturing feature interactions and mitigating overfitting. This robust performance naturally led to XGBoost, which achieved an accuracy of 0.72, precision of 0.75, recall of 0.81, F1-score of 0.78, and ROC-AUC of 0.7685, with its optimised hyperparameters derived from grid search contributing to its stability and recall performance. The ensemble models, Random Forest and XGBoost, thus outperformed KNN, LR, and SVM, achieving the highest accuracy (0.72) and F1-scores (0.78), alongside competitive ROC-AUC values (0.7689 and 0.7685, respectively), reflecting their adeptness at handling the imbalanced dataset post-SMOTE.

The high recall values across the models, particularly for LR (0.88) and XGBoost (0.81), are particularly significant, indicating a strong capability to identify true KMC cases, a critical factor in healthcare settings where missing eligible children could lead to adverse outcomes. This strength in recall links the models' practical utility, with XGBoost's slight edge over Random Forest, despite their similar metrics, likely due to its optimised hyperparameters and regularisation, which enhance its adaptability to the dataset's characteristics. Collectively, these results position ensemble methods, especially XGBoost, as highly suitable for predicting KMC adoption, laying a reliable foundation for deeper interpretability through SHAP and LIME, which can further elucidate the feature contributions and support clinical decision-making in this context.

3.3. Implications for prediction of KMC adoption

The results have significant implications for predicting which children are placed on their mother's chest and bare skin after birth, a cornerstone of KMC that promotes early bonding and reduces neonatal mortality. The high recall of XGBoost (0.81) and Random Forest (0.80) indicates that these models can effectively identify a large proportion of children eligible for KMC, enabling targeted interventions in Sierra Leone's resource-constrained healthcare system. The identified features, such as the Number of antenatal visits and the Place of delivery, suggest that enhancing prenatal care and institutional delivery infrastructure could increase KMC uptake, particularly in regions with lower baseline rates. The socio-economic factors (e.g., health insurance coverage, current employment) highlight the need for policy interventions to address financial barriers, ensuring that KMC remains accessible regardless of maternal employment or insurance status.

Table 1
Performance Evaluation of ML models for Prediction of KMC.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.61	0.70	0.60	0.65	0.66
Logistic Regression	0.68	0.68	0.88	0.77	0.6285
SVM (RBF Kernel)	0.67	0.71	0.76	0.73	0.6949
Random Forest	0.72	0.75	0.80	0.78	0.7689
XGBoost	0.72	0.75	0.81	0.78	0.7685

The regional variability captured by region and sampling variables underscores the importance of localised strategies, allowing healthcare providers to prioritise high-need areas. The strong performance of XGBoost, combined with SHAP and LIME interpretations, offers a transparent framework for clinicians to understand which features (e.g., Caesarean delivery, antenatal visits) drive predictions, facilitating informed decision-making at the point of care. However, the moderate precision (0.75 for XGBoost) suggests a risk of false positives, with some children incorrectly flagged for KMC. This trade-off is acceptable in this context, given the low-risk nature of KMC and the priority of maximising recall to ensure no eligible child is missed. Future research could explore threshold tuning or cost-sensitive learning to further balance precision and recall. Overall, these findings provide an actionable tool for healthcare policymakers and practitioners to enhance KMC adoption, potentially reducing neonatal morbidity and mortality in Sierra Leone.

3.4. LIME interpretation for prediction of KMC

LIME (Local Interpretable Model-agnostic Explanations) was employed to generate instance-level explanations of the model's predictions. Six representative cases were plotted, including both high-confidence and ambiguous outputs. Each LIME plot showed how specific features contributed positively or negatively toward the predicted class. Features such as "Covered by health insurance" and "Delivery by caesarean section" frequently contributed negatively to predictions for Class 0, thereby favouring Class 1. In contrast, "Region", "Sample strata for sample errors", and, occasionally, "Primary sampling unit" contributed positively to Class 0. The direction and strength of these contributions varied across instances, emphasising the model's decision-making's contextual nature.

For high-confidence predictions (e.g., 93 % for Class 0 in Fig. 1A), a small number of dominant features pushed heavily in one direction. For borderline cases (e.g., 56 % for Class 0 in Fig. 1B), the plots revealed an approximate balance between supporting and opposing features, indicating indecision or data ambiguity. This pattern demonstrates how LIME can reveal not just what the model relies on, but also where it may be unsure.

The interpretability of the XGBoost model's predictions for Kangaroo Mother Care (KMC) adoption is further elucidated through SHAP waterfall plots in Fig. 2, which collectively showcase multiple instances from the test set. These plots provide a detailed, instance-level decomposition of feature contributions, offering a transparent view of how individual predictions are derived and enhancing the generalisability and trustworthiness of the model for neonatal care decision-making in Sierra Leone.

SHAP (SHapley Additive exPlanations) was applied to augment LIME by providing both global feature importance and local instance-specific breakdowns. The SHAP summary plot revealed significant global predictors, notably "Place of delivery", "Sample strata for sample errors", and "Women's individual sample weight". These features showed high average SHAP values, reflecting a strong overall influence across the dataset.

SHAP waterfall plots for selected instances reinforced the interpretability provided by LIME. In confident predictions, SHAP values showed aligned contributions that firmly pushed the output toward Class 1, whereas in uncertain or misclassified instances, opposing forces from different features were nearly equal. For example, in one misclassified instance (true Class 0, predicted Class 1), SHAP clearly showed that a cluster of positive contributions overwhelmed weaker negative contributions, explaining the model's misstep.

Unlike LIME, SHAP's additive explanation framework quantitatively decomposes the model's output from a baseline prediction. This offers a more precise, probabilistic view of how each feature shifts the output, which is especially useful for understanding borderline predictions or

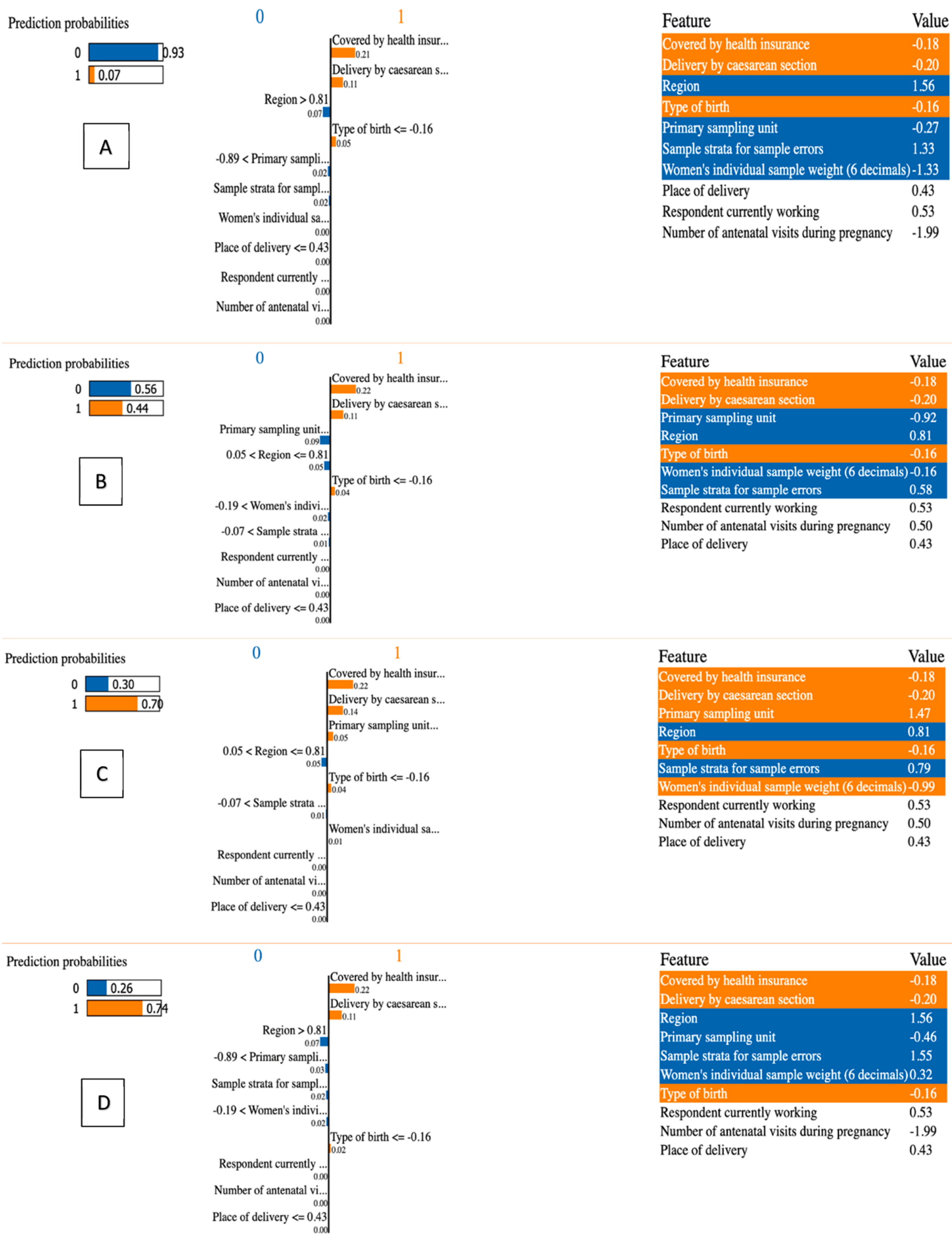

Fig. 1. LIME explanations for six individual predictions, showing key feature contributions to class 0 and class 1 decisions.

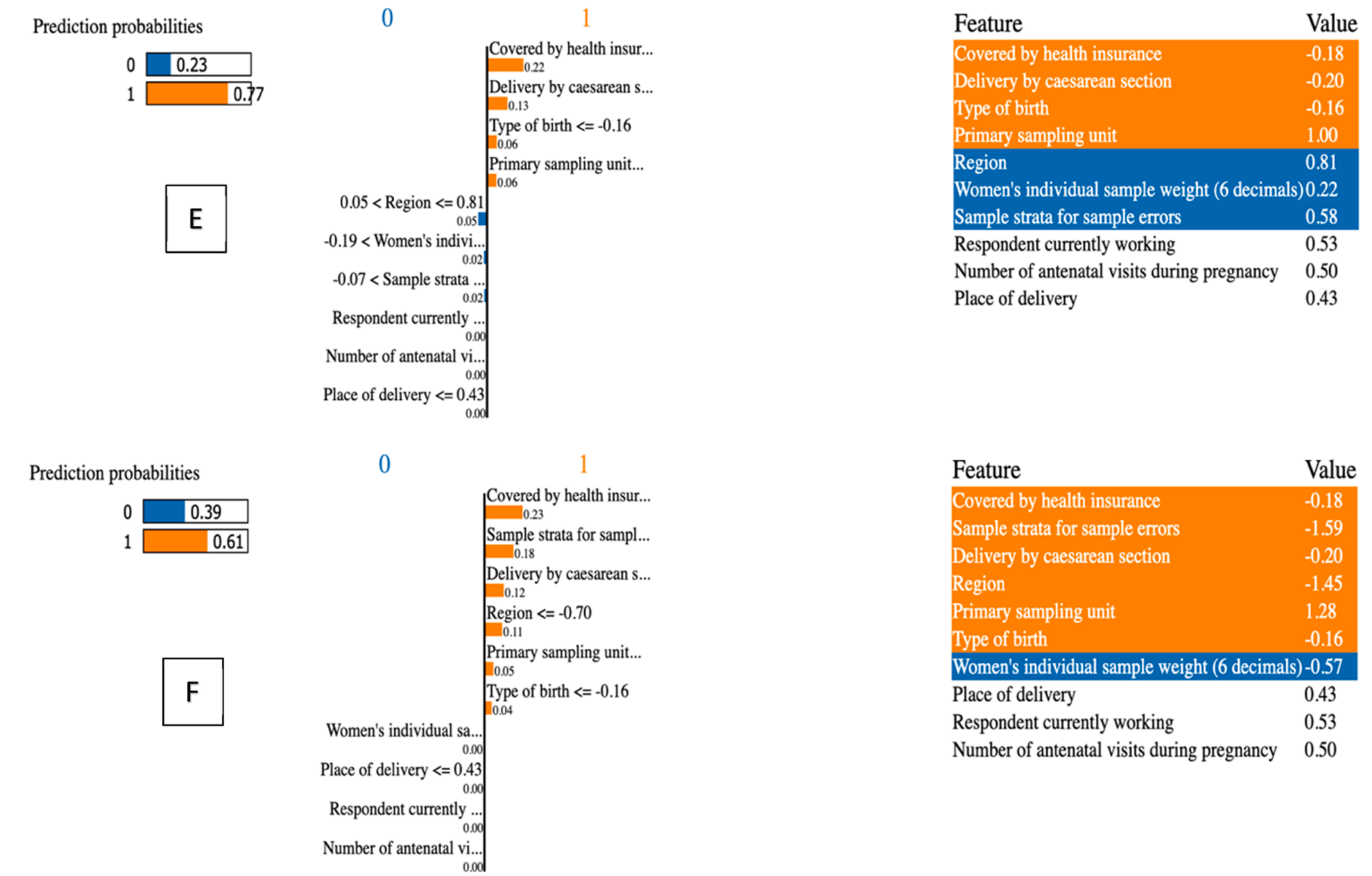

Fig. 1. (*continued*).

edge cases.

The global interpretability of the XGBoost model's predictions for Kangaroo Mother Care (KMC) adoption is comprehensively illustrated in the SHAP summary plot shown in Fig. 3. This plot provides an overview of feature importance across the dataset, ranking the 10 selected features by their mean absolute SHAP values and indicating their direction of impact, offering valuable insights into the key drivers of KMC adoption in the Sierra Leone context.

The comparative analysis of feature contributions to the XGBoost model's predictions for KMC adoption is presented in Table 2, which synthesises the insights from LIME and SHAP. Table 2 evaluates the presence and impact of the 10 selected features, spanning clinical, socioeconomic, and sampling-related variables across both methods, and assesses the degree of interpretability agreement to provide a robust framework for understanding the drivers of KMC adoption in Sierra Leone.

Whilst both LIME and SHAP aim to enhance model transparency, they do so using fundamentally different approaches. LIME creates local surrogate models based on perturbed samples, whereas SHAP relies on Shapley values from cooperative game theory, ensuring consistency and additivity.

Across the selected cases, both methods identified consistent key features, including health insurance coverage, caesarean delivery, and region. These features were repeatedly implicated in predictions for Class 1 or Class 0, regardless of the explanation method. However, SHAP tended to capture more granular variation in contribution magnitudes, whilst LIME provided intuitive visual groupings of how features contributed to specific predictions.

In terms of complementarity, LIME is particularly useful for explaining specific decisions to non-technical stakeholders, whereas SHAP is better suited for detailed audits or feature attribution analysis. Their alignment on the most impactful features enhances confidence in the interpretability outputs, whilst their differences underscore the importance of using multiple explanation tools for robust model evaluation.

4. Discussion

This study represents the first systematic application of ML techniques combined with explainable AI to predict KMC adoption in Sierra Leone, providing novel insights into the factors that influence this critical neonatal intervention. The superior performance of ensemble methods, particularly XGBoost with 72 % accuracy and 81 % recall, aligns with recent findings in maternal health prediction studies. Khadidos et al. [12] demonstrated that ensemble machine learning frameworks achieved similar performance levels when predicting maternal health risks during pregnancy, suggesting that ensemble approaches offer robust solutions for complex healthcare prediction tasks.

The effectiveness of machine learning approaches in predicting maternal and neonatal health outcomes has been consistently demonstrated across multiple domains. Rezaei et al. [33] successfully employed linear regression models to predict emergency cesarean sections among nulliparous women, achieving an impressive AUC of 0.86, accuracy of 0.82, and recall of 0.85. Their study identified advanced maternal age, diabetes, preeclampsia, and doula support as key predictors, factors that resonate with our finding that delivery characteristics and maternal health conditions significantly influence KMC adoption. Similarly, Taeidi et al. [34] demonstrated that Deep Learning models could predict intrauterine growth restriction with an AUROC of 0.91, highlighting drug addiction, previous history of intrauterine growth restriction, chronic hypertension, and preeclampsia as weighted factors. This parallels our identification of delivery method and maternal health status as critical KMC predictors.

Furthermore, Safarzadeh et al. [37] used machine learning to predict skin-to-skin contact implementation, with their Deep Learning model achieving an AUROC of 0.81, similar to our XGBoost performance. They identified doula support, neonatal weight, gestational age, and attending childbirth classes as critical predictors, reinforcing our findings on the importance of prenatal care and support systems. Roozbeh et al. [35] successfully predicted non-reassuring fetal heart patterns using Random Forest classification, achieving an AUROC of 0.77, whilst Banaei et al. [36] predicted episiotomy risk using linear regression,

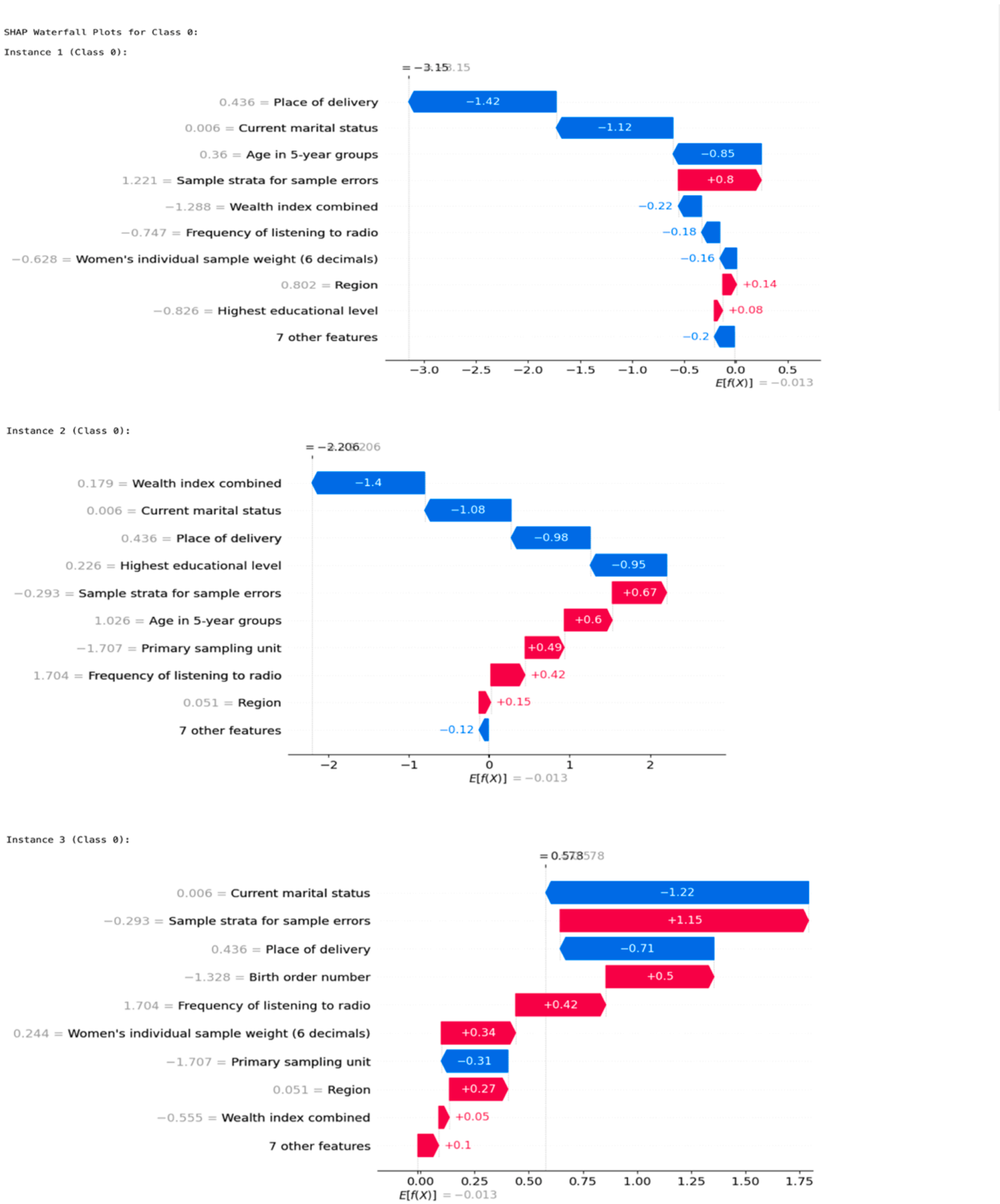

Fig. 2. SHAP summary plot and waterfall plots for selected instances, illustrating global feature importance and local prediction breakdowns.

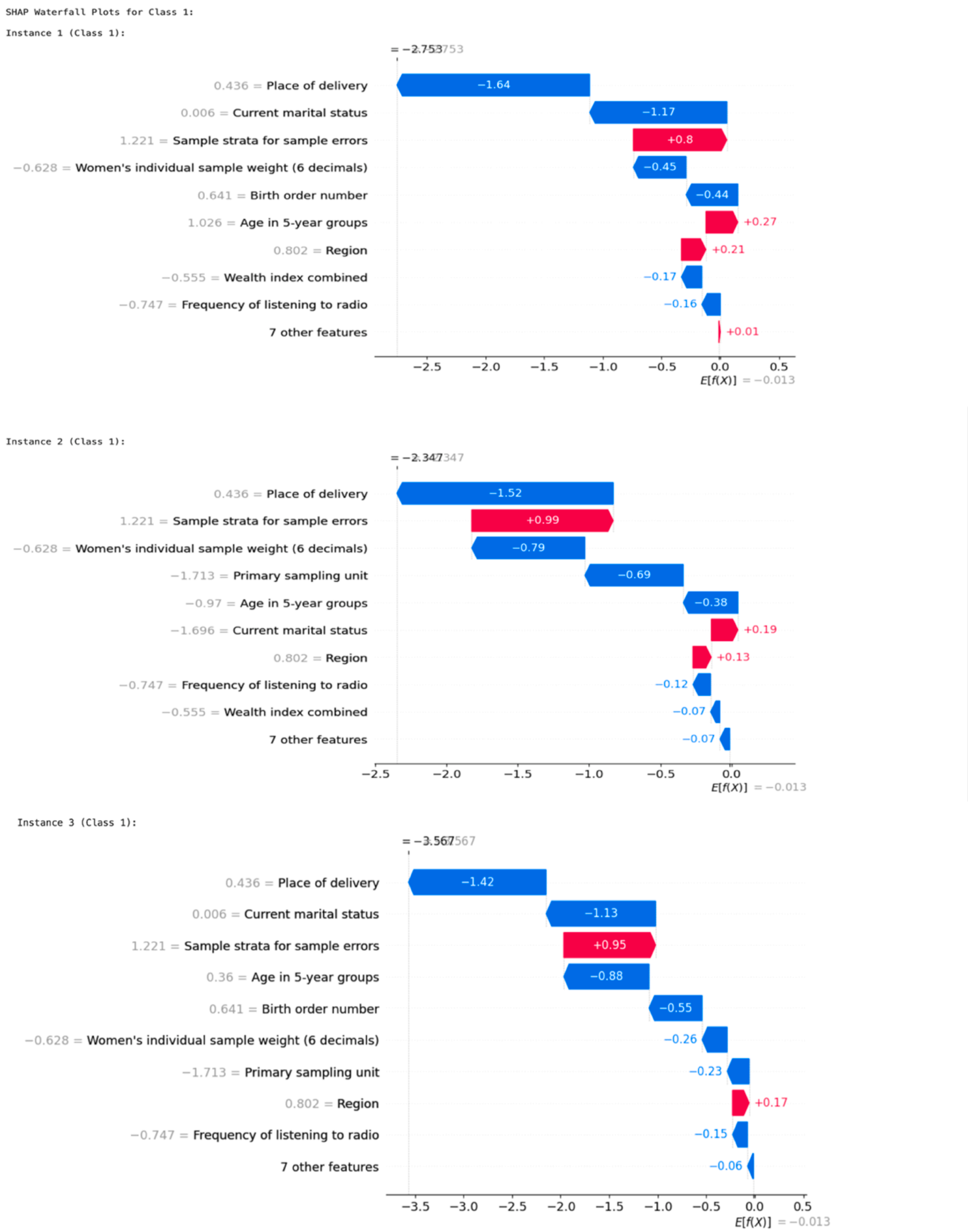

Fig. 2. (*continued*).

achieving an AUC of 0.85. These studies collectively demonstrate that ML models with AUCs between 0.70 and 0.90 are effective for clinical prediction tasks in maternal and neonatal healthcare, placing our XGBoost model's AUROC (0.7685) within the established range of acceptable clinical utility.

Predictive features identified in this study, such as delivery method, antenatal care utilisation, socio-economic status, and regional factors, are consistent with broader literature on KMC implementation barriers and enablers [21]. The recent OMWaNA trial in Uganda highlighted the importance of delivery characteristics and healthcare system factors in KMC implementation, though their focus was on clinical stabilisation rather than adoption prediction [22]. Our findings extend this understanding by quantifying the relative importance of different factors using ML approaches. Multi-country analyses have identified three pathways to KMC scale-up: champion-led, project-initiated, and health systems-designed, with socio-economic and delivery characteristics featuring prominently across all pathways [6]. Socio-economic status is particularly profound, as women's empowerment strategies have been identified as a key determinant of neonatal mortality [23].

The application of SHAP and LIME explainability techniques in our study addresses a critical gap in healthcare AI applications, where model transparency is essential for clinical acceptance. Recent systematic

Fig. 3. SHAP summary plot.

Table 2
Comparison of LIME and SHAP interpretability outputs across shared features.

Feature	LIME Presence	SHAP Presence	Interpretation Agreement
Covered by health insurance	Frequent (Class 1 driver)	High impact (Class 1)	Yes
Delivery by caesarean section	Frequent (Class 1 driver)	High impact (Class 1)	Yes
Region	Frequent (Class 0 driver)	High impact (Class 0)	Yes
Primary sampling unit	Frequent (mixed)	High impact (mixed)	Partial
Sample strata for sample errors	Frequent (mixed)	High impact (mixed)	Partial
Women's individual sample weight	Frequent (mixed)	Moderate, instance-dependent	Partial
Place of delivery	Occasional	Top contributor	Yes
Respondent currently working	Occasional	Low but consistent	Yes
Number of antenatal visits	Occasional	Low, variable	Yes
Type of birth	Consistent, small effect	Low impact	Yes

reviews emphasise that SHAP and LIME frameworks have become popular interpretive tools for healthcare applications, providing both local and global insights that enhance model transparency and trust [18]. Our comparative analysis, which shows agreement between LIME and SHAP on key features (health insurance, caesarean delivery, regional factors), strengthens confidence in the interpretability of the outputs. However, Salih et al. [20] showed that SHAP and LIME can be affected by adopted ML models and feature collinearity, necessitating careful interpretation of their outputs, which supports our decision to use both methods complementarily.

The moderate precision (75 %) and high recall (81 %) in our XGBoost model reflect an essential trade-off in healthcare prediction tasks. Similar patterns have been observed in other pregnancy outcome prediction studies, where achieving high sensitivity for identifying at-risk cases often comes at the cost of increased false positives [10]. Henry [24] found that predicting newborn birth outcomes with ML achieved comparable performance metrics, with precision-recall trade-offs being common in maternal health applications. In the context of KMC prediction, this trade-off is clinically acceptable given the low-risk nature of the intervention and the critical importance of identifying all eligible infants.

The regional disparities captured in our model align with known challenges in Sierra Leone's healthcare system. Recent studies have identified significant barriers to KMC implementation in sub-Saharan Africa, including healthcare provider knowledge gaps, infrastructure limitations, and challenges in policy implementation [21,25]. Tumukunde et al. [26] found that implementation barriers and facilitators for KMC in Uganda varied significantly by region and healthcare facility type, supporting our finding that place of delivery and regional factors are strong predictors. The WHO's updated Global Position Paper and Implementation Strategy [27] emphasises the need for health facility administrators and programme managers to focus on system change for KMC implementation, directly supporting our model's emphasis on institutional and policy factors.

Compared to traditional risk assessment approaches, our ML framework offers several advantages. Mapari et al. [11] highlighted AI's potential to revolutionise maternal health by enhancing care and accessibility, particularly in resource-constrained settings where clinical expertise may be limited. Hossain et al. [14] demonstrated that medical cyber-physical systems using machine learning could effectively predict maternal health risks in developing countries, achieving accuracy comparable to that of our study. Our results corroborate these findings whilst extending them to the specific context of KMC adoption prediction.

The practical implications of our findings extend beyond prediction accuracy to inform targeted intervention strategies. Togunwa et al. [13] emphasised that machine learning applications in maternal and fetal health must be translated into actionable clinical insights. Our model could be integrated into existing healthcare information systems to provide real-time KMC adoption risk assessment. The explainable AI component addresses the critical need for transparency in clinical decision support systems, as healthcare professionals require AI applications to be transparent about their decision-making processes to gain trust.

Recent work by Islam et al. [10] in their systematic review of machine learning for pregnancy outcome prediction found that studies achieving 70–82 % classification accuracy were considered satisfactory for clinical decision support, positioning our results within established performance benchmarks. However, they also noted that most studies lacked external validation and implementation assessment, highlighting important directions for future research.

The ensemble methods' superior performance in our study is consistent with recent comparative analyses. Mashrafi et al. [28] found that ensemble methods such as Random Forest and XGBoost consistently outperformed traditional statistical methods when predicting maternal risk levels using nationwide datasets. Their emphasis on the importance of feature selection and model interpretability aligns with our methodological approach. Recent systematic reviews have also highlighted that ensemble methods achieve superior performance in healthcare prediction tasks compared to individual algorithms [10].

However, our study's focus on Sierra Leone limits direct generalisability to other contexts, though the methodological framework could be adapted for different settings. Recent narrative reviews emphasise that ML models in maternal and fetal health often require adaptation to local healthcare contexts and populations [10,13]. Future research should explore the transferability of our approach to other Sub-Saharan African countries with similar healthcare challenges, whilst accounting for context-specific factors.

4.1. Strategies to enhance model accuracy

While the XGBoost model achieved a satisfactory accuracy of 72 % and a high recall of 81 %, there is room to further improve its predictive performance. Potential strategies include threshold calibration to optimize the balance between precision and recall, addressing the moderate precision (75 %) observed in our results. Incorporating additional features not captured in the current dataset, such as cultural beliefs, family support systems, and healthcare provider attitudes, could provide a more comprehensive understanding of KMC adoption factors. Furthermore, leveraging more advanced ensemble techniques or deep learning models, if computational resources permit, may capture complex patterns in larger or more diverse datasets. Increasing the dataset size through multi-center studies or collaborations could also improve model generalization. Finally, cost-sensitive learning approaches could be explored to further address the precision-recall trade-off, prioritizing the identification of true KMC cases in resource-constrained settings.

The internal and external validity of our findings are critical considerations for their practical application. Internally, the robustness of our methodology is supported by stratified train-test splitting, SMOTE for class imbalance, hyperparameter optimization via grid search, and the use of multiple performance metrics (accuracy, precision, recall, F1-score, ROC-AUC) to evaluate model performance comprehensively. These steps ensure that the model captures underlying data patterns reliably. Externally, the generalizability of our results is limited by the focus on Sierra Leone's healthcare system and facility-based deliveries, which may not fully represent rural or home births. Future research should prioritize external validation in diverse Sub-Saharan African contexts to assess the transferability of our predictive framework, adapting it to local healthcare systems and populations as needed.

4.2. Limitations

Several limitations must be acknowledged in interpreting these findings. First, the cross-sectional nature of the data limits our ability to establish causal relationships between predictive features and KMC adoption. Whilst machine learning models can identify associations and patterns, they cannot definitively prove that interventions targeting identified factors will improve KMC uptake without prospective validation studies.

Second, the dataset size, whilst adequate for machine learning applications, may limit the generalisability of findings to the broader Sierra Leonean population. The sample may not fully represent rural or hard-to-reach populations where KMC adoption patterns could differ significantly. Additionally, the dataset's focus on facility-based deliveries may introduce selection bias, as home births and deliveries in smaller health facilities are potentially underrepresented.

Third, the feature selection process, whilst systematic, may have excluded relevant variables that could influence KMC adoption. Important factors such as cultural beliefs, family support systems, healthcare provider attitudes, and facility-specific policies were not captured in the available dataset. These unmeasured confounders could influence both the predictive accuracy and the interpretability of the models.

Fourth, the temporal context of data collection may affect the relevance of findings to current practice. Healthcare systems, policies, and KMC implementation strategies evolve, and models trained on historical data may not accurately reflect current adoption patterns.

Fifth, the explainability techniques (SHAP and LIME), whilst providing valuable insights, have known limitations including sensitivity to model choice and feature correlations. The interpretations should be considered as approximate explanations rather than definitive causal explanations of model behaviour.

Sixth, the SMOTE technique used to address class imbalance, whilst improving model performance on minority classes, may introduce synthetic data points that do not reflect real-world scenarios. This could affect the model's performance when applied to new, unseen data.

Additionally, the temporal context of data collection may affect the relevance of findings to current practice, as healthcare systems and KMC implementation strategies evolve. Models trained on historical data may not fully reflect contemporary adoption patterns. Finally, while our model identifies factors associated with KMC adoption, it does not provide evidence that targeting these factors through specific interventions will improve outcomes. Prospective implementation studies are needed to validate the clinical utility of the predictive framework and ensure its alignment with real-world healthcare needs.

Finally, the study's focus on prediction rather than intervention limits its immediate clinical utility. Whilst the model identifies factors associated with KMC adoption, it does not provide evidence that targeting these factors through specific interventions will improve outcomes. Prospective implementation studies would be needed to validate the clinical utility of the predictive framework.

5. Conclusion

This study demonstrates that combining machine learning with explainable AI effectively predicts KMC adoption in Sierra Leone. We employed five ML classifiers (KNN, LR, SVM, RF, and XGBoost) on a nationally representative dataset of 7737 births, using forward-backward feature selection to identify 10 key predictors from 22 candidate variables. XGBoost achieved the strongest performance with 72 % accuracy, 75 % precision, 81 % recall, an F1-score of 0.78, and an ROC-AUC of 0.7685. SHAP and LIME explainability analyses consistently identified delivery characteristics (caesarean section, place of delivery), socio-economic factors (maternal employment, health insurance coverage), and prenatal care utilization (antenatal visits) as the primary drivers of KMC adoption.

The high recall (81 %) makes this model particularly suitable for screening applications in resource-constrained healthcare settings, ensuring that most eligible infants are identified for KMC counselling and support. The identified modifiable factors, especially antenatal care attendance, health insurance access, and institutional delivery, provide actionable targets for policy interventions aimed at improving KMC uptake rates. The explainable AI framework enhances clinical acceptance by making model predictions transparent and interpretable to healthcare providers and policymakers.

We recommend prospective validation of this predictive model in clinical settings, integration into existing health information systems for real-time decision support, and the development of targeted interventions to address the identified predictors. Future research should explore threshold calibration to optimise the precision-recall balance, assess the model's transferability to other Sub-Saharan African contexts, and evaluate the impact of model-guided interventions on actual KMC adoption rates and neonatal health outcomes.

Funding sources

No funding was received for this study.

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

NA

Data availability

The data used for this study is freely available at https://dhsprogram.com/data/dataset/Sierra-Leone_Standard-DHS_2019.cfm?flag=0.

CRediT authorship contribution statement

Afeez A. Soladoye: Writing – review & editing, Writing – original draft, Methodology, Investigation, Software, Formal analysis, Conceptualization. **David B. Olawade:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis. **Joseph E. Origbo:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation. **Kobloobase O. Usani:** Writing – review & editing, Writing – original draft, Methodology, Data curation, Investigation. **Ayomide Israel Adekoya:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Ojima Z. Wada:** Writing – review & editing, Investigation. **Augustus Osborne:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Software.

Declaration of competing interest

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

Acknowledgements

NA

References

[1] E.O. Boundy, R. Dastjerdi, D. Spiegelman, W.W. Fawzi, S.A. Missmer, E. Lieberman, et al., Kangaroo mother care and neonatal outcomes: a meta-analysis, Pediatrics. 137 (2016), https://doi.org/10.1542/peds.2015-2238.

[2] A. Conde-Agudelo, D.R. JL, Kangaroo mother care to reduce morbidity and mortality in low birthweight infants, Cochrane Database Syst. Rev. 2016 (2016) CD002771, https://doi.org/10.1002/14651858.CD002771.pub4.

[3] WHO, Launch of Kangaroo Mother Care Global position paper and Implementation strategy, Newsroom, https://www.who.int/news-room/events/detail/2023/05/16/default-calendar/launch-of-kangaroo-mother-care-global-position-paper-and-implementation-strategy, 2023 (accessed August 27, 2025).

[4] E.R. Smith, I. Bergelson, S. Constantian, B. Valsangkar, G.J. Chan, Barriers and enablers of health system adoption of Kangaroo mother care: a systematic review of caregiver perspectives, BMC. Pediatr. 17 (2017) 35, https://doi.org/10.1186/s12887-016-0769-5.

[5] Q. Cai, Y. Zhou, M. Hong, D. Chen, X. Xu, Healthcare providers' perceptions and experiences of kangaroo mother care for preterm infants in four neonatal intensive care units in China: a qualitative descriptive study, Front. Public Health 12 (2024), https://doi.org/10.3389/fpubh.2024.1419828.

[6] L. Vesel, A.-M. Bergh, K.J. Kerber, B. Valsangkar, G. Mazia, S.G. Moxon, et al., Kangaroo mother care: a multi-country analysis of health system bottlenecks and potential solutions, BMC. Pregnancy. ChildBirth 15 (2015), https://doi.org/10.1186/1471-2393-15-S2-S5. S5.

[7] W.Z. Ojima, D.B. Olawade, O.O. Awe, A.O. Amusa, Factors associated with neonatal mortality among newborns admitted in the special care baby unit of a Nigerian hospital, J. Trop. Pediatr. 67 (2021), https://doi.org/10.1093/tropej/fmab060.

[8] D.B. Olawade, O.Z. Wada, N. Aderinto, A. Odetayo, Y.A. Adebisi, D.T. Esan, et al., Factors contributing to under-5 child mortality in Nigeria: a narrative review, Medicine 104 (2025) e41142, https://doi.org/10.1097/MD.0000000000041142.

[9] U. Sesay, A. Osborne, Building climate-resilient health systems in Sierra Leone: addressing the dual burden of infectious and climate-related diseases, Infect. Dis. Poverty. 14 (2025) 23, https://doi.org/10.1186/s40249-025-01294-9.

[10] M.N. Islam, S.N. Mustafina, T. Mahmud, N.I. Khan, Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda, BMC. Pregnancy ChildBirth 22 (2022) 348, https://doi.org/10.1186/s12884-022-04594-2.

[11] S.A. Mapari, D. Shrivastava, A. Dave, G.N. Bedi, A. Gupta, P. Sachani, et al., Revolutionizing maternal health: the role of artificial intelligence in enhancing care and accessibility, Cureus (2024), https://doi.org/10.7759/cureus.69555.

[12] A.O. Khadidos, F. Saleem, S. Selvarajan, Z. Ullah, A.O. Khadidos, Ensemble machine learning framework for predicting maternal health risk during pregnancy, Sci. Rep. 14 (2024) 21483, https://doi.org/10.1038/s41598-024-71934-x.

[13] T.O. Togunwa, A.O. Babatunde, K.-R. Abdullah, Deep hybrid model for maternal health risk classification in pregnancy: synergy of ANN and random forest, Front. Artif. Intell. 6 (2023), https://doi.org/10.3389/frai.2023.1213436.

[14] M.M. Hossain, M.A. Kashem, N.M. Nayan, M.A. Chowdhury, A medical cyber-physical system for predicting maternal health in developing countries using machine learning, Healthc. Anal. 5 (2024) 100285, https://doi.org/10.1016/j.health.2023.100285.

[15] Z. Sadeghi, R. Alizadehsani, C.I.F.C.I. MA, S. Kausar, R. Rehman, P. Mahanta, et al., A review of explainable artificial Intelligence in healthcare, Comput. Electr. Eng. 118 (2024) 109370, https://doi.org/10.1016/j.compeleceng.2024.109370.

[16] Ribeiro M.T., Singh S., Guestrin C. "Why should I trust you?": explaining the predictions of any classifier. ArXiv 2016. https://doi.org/10.48550/arXiv.1602.04938.

[17] Lundberg S., Lee S.-I. A unified approach to interpreting model predictions. ArXiv 2017. https://doi.org/10.48550/arXiv.1705.07874.

[18] V. Vimbi, N. Shaffi, M. Mahmud, Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection, Brain Inf. 11 (2024) 10, https://doi.org/10.1186/s40708-024-00222-1.

[19] A. Sathyan, A.I. Weinberg, K. Cohen, Interpretable AI for bio-medical applications, Complex Eng. Syst. 2 (2022) 18, https://doi.org/10.20517/ces.2022.41.

[20] A.M. Salih, Z. Raisi-Estabragh, I.B. Galazzo, P. Radeva, S.E. Petersen, K. Lekadir, et al., A perspective on explainable artificial intelligence methods: SHAP and LIME, Adv. Intell. Syst. 7 (2025), https://doi.org/10.1002/aisy.202400304.

[21] S.M. Bilal, H. Tadele, T.A. Abebo, B.T. Tadesse, M. Muleta, F. W/Gebriel, et al., Barriers for kangaroo mother care (KMC) acceptance, and practices in southern Ethiopia: a model for scaling up uptake and adherence using qualitative study, BMC. Pregnancy. ChildBirth 21 (2021) 25, https://doi.org/10.1186/s12884-020-03409-6.

[22] M.M. Medvedev, V. Tumukunde, I. Mambule, C.J. Tann, P. Waiswa, R.R. Canter, et al., Operationalising kangaroo mother care before stabilisation amongst low birth weight neonates in Africa (OMWaNA): protocol for a randomised controlled trial to examine mortality impact in Uganda, Trials 21 (2020) 126, https://doi.org/10.1186/s13063-019-4044-6.

[23] J.-M.V. Idowu, W.Z. Ojima, S.B. Adetutu, M.A Mayowa, A. Joseph Oluwakayode, O. Tubosun Alex, Women's empowerment as a determinant of neonatal mortality in Sub-Saharan Africa: a narrative review focused on Nigeria, Glob. Health Action. 17 (2024), https://doi.org/10.1080/16549716.2024.2394256.

[24] R.D. Henry, Predicting newborn birth outcomes with prenatal maternal health features and correlates in the United States: a machine learning approach using archival data, BMC Pregnancy. ChildBirth 24 (2024) 603, https://doi.org/10.1186/s12884-024-06812-5.

[25] B. Afande Mukhola, L.W. Kivuti-Bitok, A. Chepchirchir, Challenges faced by mothers practicing kangaroo mother care (KMC) in a resource-limited setting during the COVID-19 pandemic: insights from the voices of mothers of preterm babies and their suggested solutions, Glob. Pediatr. Health 10 (2023), https://doi.org/10.1177/2333794X231213472.

[26] V.S. Tumukunde, J. Katongole, S. Namukwaya, M.M. Medvedev, M. Nyirenda, C. J. Tann, et al., Kangaroo mother care prior to clinical stabilisation: implementation barriers and facilitators reported by caregivers and healthcare providers in Uganda,

PLOS. Glob. Public Health 4 (2024) e0002856, https://doi.org/10.1371/journal.pgph.0002856.
[27] WHO. Kangaroo mother care - implementation strategy for scale-up adaptable to different country contexts. 2023.
[28] S.S. Al Mashrafi, L. Tafakori, M. Abdollahian, Predicting maternal risk level using machine learning models, BMC Pregnancy. ChildBirth 24 (2024) 820, https://doi.org/10.1186/s12884-024-07030-9.
[29] Nelson M., Rajendran S. Detection of neonatal jaundice using capsule graph neural network optimized with wild horse algorithm. InIET Conference Proceedings CP913 2024 Dec 1 (Vol. 2024, No. 37, pp. 787–791). Stevenage, UK: The Institution of Engineering and Technology.
[30] M. Nelson, S. Rajendran, O.I. Khalaf, H. Hamam, Deep-learning-based intelligent neonatal seizure identification using spatial and spectral GNN optimized with the Aquila algorithm, AIMS Math. 9 (7) (2024) 19645–19669. Jan 1.
[31] M. Nelson, S. Rajendran, Y. Alotaibi, Vision graph neural network-based neonatal identification to avoid swapping and abduction, AIMS Math. 8 (2023) 21554–21571.
[32] N. Madhusundar, R. Surendran, Neonatal jaundice identification over the face and sclera using graph neural networks, in: In2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), 2023, pp. 1243–1249. Jan 23IEEE.
[33] N. Rezaei, M. Amani, H. Asgharpoor, V. Mehrnoush, F. Darsareh, A. Nikmanesh, Machine learning approach to predict emergency cesarean sections among nulliparous women, Cureus. 17 (8) (2025). Aug 19.
[34] E. Taeidi, A. Ranjbar, F. Montazeri, V. Mehrnoush, F. Darsareh, Machine learning-based approach to predict intrauterine growth restriction, Cureus. 15 (7) (2023). Jul 6.
[35] N. Roozbeh, F. Montazeri, M.V. Farashah, V. Mehrnoush, F. Darsareh, Proposing a machine learning-based model for predicting nonreassuring fetal heart, Sci. Rep. 15 (1) (2025) 7812. Mar 6.
[36] M. Banaei, N. Roozbeh, F. Darsareh, V. Mehrnoush, M.S. Farashah, F. Montazeri, Utilizing machine learning to predict the risk factors of episiotomy in parturient women, AJOG. Glob. Rep. 5 (1) (2025) 100420. Feb 1.
[37] S. Safarzadeh, N.S. Ardabili, M.V. Farashah, N. Roozbeh, F. Darsareh, Predicting mother and newborn skin-to-skin contact using a machine learning approach, BMC. Pregnancy. ChildBirth 25 (1) (2025) 182. Feb 18.