RESEARCH ARTICLE

Harnessing Digital Health and Machine Learning for Early Detection of Hypertensive Disorders in Pregnancy: Insights from low-resource settings, Tanzania

[version 1]

Isaac Lyatuu[1,2], Emmanuel Mwanga[2], Yusuph Kulindwa[1], Raymond Bandio[1], Alen Kinyina[1], Ntuli Kapologwe[3], Phineas Sospeter[3], Ahmad Makuwani[3], James Tumaini[4,5], Omari Sukari[6], Augustino Hellar[1]

[1]Prime Health Initiative Tanzania, Dar-es-Salaam, Tanzania, Tanzania
[2]Ifakara Health Institute, Ifakara, Morogoro Region, Tanzania
[3]Ministry of Health Tanzania, Dodoma, Tanzania
[4]President's Office Regional Authority and Local Government, Dodoma, Tanzania
[5]The University of Dodoma College of Health Sciences, Dodoma, Dodoma Region, Tanzania
[6]Regional Health Secretariat, Geita Region, Geita, Tanzania

v1 **First published:** 30 Jun 2025, **2**:158
https://doi.org/10.12688/verixiv.1245.1

Latest published: 30 Jun 2025, **2**:158
https://doi.org/10.12688/verixiv.1245.1

Abstract

Introduction

Tanzania's high maternal mortality ratio presents economic and social challenges, particularly in marginalized communities. Key contributors include inadequate antenatal care (ANC), healthcare provider shortages, and limited diagnostic resources. Digital health and machine learning (ML) offer new opportunities for early risk identification. This study explores the potential of leveraging digital ANC records to predict hypertensive disorders during pregnancy in low-resource settings.

Methods

We analyzed 300,000 ANC visit records from the Unified Community System (UCS) (2023–2024), a Government of Tanzania-led digital data collection initiative. The dataset included key maternal health indicators such as blood pressure, BMI, and urine protein levels.

Hypertensive disorder was defined as systolic BP ≥140 mmHg and/or diastolic BP ≥90 mmHg. Five ML classifiers—K-Nearest Neighbors, Logistic Regression, Random Forest, Support Vector Machine, and XGBoost—were tested. The final model was trained on 187,438 records and validated using 120,000 independent records.

Results

A total of 1,894 cases met the outcome criteria. The training dataset consisted of 3,788 records (50/50 risk distribution). XGBoost was selected as the final model, achieving 90.05% accuracy, 14% precision, and a 24% F1-score for gestational hypertension prediction. The AUC was 0.95 (95% CI: 0.9532–0.9548), indicating strong discriminatory power. The model was deployed online for real-world application.

Conclusion

AI-driven risk prediction models can enhance early detection of hypertensive disorders, reduce maternal mortality, and ease healthcare burdens. However, improving routine data quality is essential for scalability and reliability. Policymakers should prioritize AI adoption alongside robust data and implementation strategies.

Keywords
Machine learning, maternal health, pregnancy complications, hypertensive disorders, pre-eclampsia, eclampsia

Corresponding author: Isaac Lyatuu (ilyatuu@ihi.or.tz)

Author roles: Lyatuu I: Conceptualization, Formal Analysis, Methodology, Software, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Mwanga E**: Formal Analysis, Software, Visualization; **Kulindwa Y**: Funding Acquisition, Writing – Review & Editing; **Bandio R**: Funding Acquisition, Writing – Review & Editing; **Kinyina A**: Data Curation, Validation, Writing – Review & Editing; **Kapologwe N**: Funding Acquisition, Writing – Review & Editing; **Sospeter P**: Supervision, Writing – Review & Editing; **Makuwani A**: Supervision, Writing – Review & Editing; **Tumaini J**: Supervision, Writing – Review & Editing; **Sukari O**: Supervision, Writing – Review & Editing; **Hellar A**: Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project was funded by the Gates Foundation (BMGF) under the Grand Challenges initiative (Investment ID: INV-046249). Data collection and personnel support were provided through the Unified Community System, which operates as part of the Government of Tanzania's routine services under the Ministry of Health and the President's Office – Regional Administration and Local Government

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Lyatuu I, Mwanga E, Kulindwa Y *et al.* **Harnessing Digital Health and Machine Learning for Early Detection of Hypertensive Disorders in Pregnancy: Insights from low-resource settings, Tanzania [version 1]** VeriXiv 2025, **2**:158 https://doi.org/10.12688/verixiv.1245.1

First published: 30 Jun 2025, **2**:158 https://doi.org/10.12688/verixiv.1245.1

Introduction

The application of Machine Learning (ML) techniques in predicting pregnancy complications has gained significant attention in recent studies. Various algorithms, including Logistic Regression (LR), Random Forest (RF), Support Vector Machines (SVM) and Extreme Gradient Boosting (XGBoost), have been employed to predict adverse maternal outcomes for conditions such as Hypertensive Disorders of Pregnancy (HDP) (e.g., pre-eclampsia), gestational diabetes mellitus (GDM), and preterm labor.[1,2] These models utilize information such as demographics, medical images, patient histories, and clinical measurements, to identify high-risk pregnancies and guide timely intervention strategies. By leveraging the increasing availability of digital health data, recent systematic reviews highlight a parallel rise in the use of ML models to predict health outcomes, particularly through electronic medical records, medical imaging, and sensor-based data.[1,3–5] ML models excel at analyzing large, complex datasets to uncover patterns and predict complications that may be overlooked by traditional methods or constrained by operational challenges. These predictive capabilities encompass a broad range of maternal health outcomes, including severe maternal morbidities and mortalities, providing critical insights to inform preventive measures. In alignment with WHO recommendations advocating early detection and initiation of care,[6] ML-driven insights are particularly crucial in low- and middle-income countries (LMICs), where healthcare systems often face resource constraints and rely on passive reporting mechanisms.[7] Harnessing ML in such settings has the potential to enhance maternal health outcomes and address critical gaps in care delivery.

Study objectives

This study investigates the application of digital health and machine learning (ML) technologies in low-resource settings to predict the risk of HDP using routine health facility service data. The specific objectives are:

1. **To extract and analyze health records from routine antenatal care (ANC) visits** in low-resource settings, utilizing data from the Unified Community System (UCS).

2. **To develop a machine learning-based risk prediction model** for early detection of HDP during pregnancy, leveraging key maternal health indicators such as blood pressure, BMI, and urine analysis results.

3. **To test and validate the ML model's performance** within ANC routine services, assessing its predictive accuracy, precision, and applicability in real-world clinical settings.

ANC reporting

Antenatal Care (ANC) is a cornerstone of maternal health, providing critical opportunities for monitoring pregnancy progress, delivering essential preventive services, and implementing timely interventions when risks are identified. ANC services are typically delivered through visits conducted at healthcare facilities or occasionally at home by trained healthcare providers. In low- and middle-income countries (LMICs), the majority of ANC visits occur in health facilities. The World Health Organization (WHO) recommends a minimum of eight ANC visits during pregnancy to ensure comprehensive monitoring and care.[8] However, in many LMICs, a significant proportion of expectant mothers fail to attend even half of the recommended visits,[9,10] impeding efforts to effectively monitor maternal complications and contributing to persistently high maternal and neonatal mortality rates.

During ANC visits, key information is collected, including maternal age, gestational age, weight, blood pressure, urine analysis results, and fetal development indicators. This data is critical for guiding care during pregnancy and for future clinical decision-making. In many countries, particularly LMICs, this information is predominantly recorded on paper using national guidelines adapted from WHO recommendations. While paper-based systems are cost-effective and widely utilized in resource-constrained settings, they present several limitations. These include challenges in long-term data storage, difficulties in linking records across visits or facilities, and the inability to efficiently triangulate data for population-level insights. Transitioning to digital data capture systems may address these limitations, but their adoption is constrained by resource availability and infrastructure challenges in LMICs.

Innovation in ANC recording

With advancements in technology and the increasing demand to improve data quality and availability, the Government of Tanzania (GoT) has developed the Unified Community System (UCS), a software solution designed to facilitate the collection of digital health records in remote and resource-limited settings.[11] UCS is a digital platform designed to empower healthcare workers and transform community health service delivery at the grassroots level.[12] UCS operates on a client-server architecture, enabling data collection at both health facilities and community levels using tablet devices. The collected data can be temporarily stored on the device or synchronized directly to a central repository when an internet connection is available. The current Tanzania UCS design incorporates several key modules, including Client Registry, Community-Based Client Services, Antenatal Care (ANC), Labor and Delivery, Prevention of Mother-to-Child

Transmission (PMTCT), Tuberculosis and Leprosy Services and Referral System. UCS collects individual-level data and is designed in alignment with national implementation guidelines to ensure consistency with health service delivery standards. The UCS system prioritizes data security, employing password protection and encryption mechanisms to safeguard sensitive information. By addressing challenges in data collection and management, UCS represents a significant step toward enhancing health service delivery and data-driven decision-making in Tanzania.

Application of Machine Learning (ML) for predicting maternal complications

Machine learning (ML), a subset of artificial intelligence, focuses on developing computer systems capable of analyzing data, identifying patterns, and leveraging those insights to make predictions on unseen or future data. ML employs a range of predefined algorithms to uncover relationships between predictors and outcomes, enabling the creation of predictive models. In the context of maternal health, numerous studies have successfully developed and validated ML models to predict adverse maternal outcomes, including pre-eclampsia, gestational diabetes, and preterm labor. These models utilize diverse datasets, such as demographic information, medical histories, and clinical measurements, to identify at-risk pregnancies early and facilitate timely interventions.

The ML process involves training computer algorithms on existing datasets to learn classifications and patterns on and between input variables (predictors) and outcomes of interest. The trained models are then evaluated for accuracy using a test dataset, often a reserved subset of the original training data. The classification algorithm with the highest predictive accuracy is typically selected and further refined to develop the final predictive model.

Several ML classification methods can be employed depending on the nature of the input data and the specific prediction objectives. For this project, we explored and compared five widely used classification methods namely **K-Nearest Neighbor (KNN), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF)** and **Extreme Gradient Boosting (XGBoost)**. The most accurate model was identified by comparing accuracy scores and was used to serve as the foundation for predictive analysis. These approaches highlight ML's potential to enhance maternal health outcomes by enabling data-driven decision-making and targeted interventions.

Methods

Study design

This study employs a retrospective cohort design using routine antenatal care (ANC) data collected through the Unified Community System (UCS), a government-led digital health platform in Tanzania. The study focuses on developing and testing a machine learning (ML) model to predict HDP based on maternal health indicators recorded during ANC visits.

Study participants & settings

Study participants were all expectant mothers attending ANC clinics in Tanzania and were recorded in the UCS application between 2020 to 2024.

Data Extraction, Loading and Transfer (ETL)

Data was retrieved from the UCS online portal managed by the Ministry of Health. The UCS ANC dashboard contains records of first and subsequent visits as outlined in the national ANC guidelines. The dataset included all ANC visit records from the start of the UCS system through March 31, 2024. Data was downloaded in batches based on the date created. All data files were stored in a comma-separated values (CSV) format for further processing and appending. The data did not contain any person identifiable information.

The data was loaded into a Pandas data frame and a series of cleaning procedures were applied to prepare the data for ML consumption. Numerical variables erroneously stored as strings were converted into numeric formats, ensuring data consistency. Duplicate records were identified and removed using a composite variable generated from the event ID and ANC visit number. The event ID, a unique identifier created by the UCS for each recorded incident, played a critical role in maintaining the dataset's integrity.

Feature selection was guided by the outcome of interest, the availability of data points, and domain knowledge from subject matter experts, which included obstetricians and public health specialists. Data was transformed from a longitudinal structure, where multiple observations existed per client, into a cross-sectional format with one observation per client. This transformation involved aggregating key variables: gestational age was represented by its maximum value, while glucose in urine, protein in urine, syphilis status, and visit number were represented by their most recent values. Variables such as blood glucose levels, height, weight, body mass index (BMI), systolic and diastolic blood pressure, and temperature were aggregated using their mean values. Blood group, being invariant, was retained as a single recorded value. These preprocessing steps ensured the data was cleaned, structured, and appropriately formatted for the subsequent development of predictive models, enhancing its reliability and utility for scientific analysis.

Definition of the outcome of interest

The outcome of interest selected for this study was HDP, defined as a systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg. Ideally, HDP status would be directly recorded in the dataset based on clinical observations during antenatal visits. However, our analysis revealed a significant underreporting of HDP, despite evidence suggesting its prevalence in the population. Furthermore, in cases where HDP was documented, inconsistencies in data linkage prevented reliable identification of the corresponding primary clinical records.

To address these data limitations, we computed HDP arithmetically using recorded systolic and diastolic blood pressure measurements. Records with systolic or diastolic blood pressure values meeting or exceeding the thresholds of 140 mmHg and 90 mmHg, respectively, were classified as "At Risk of HDP." Conversely, records with values below these thresholds were categorized as "No Risk." This approach ensured a consistent and objective identification of the outcome variable, enabling the development of predictive models despite limitations in direct clinical reporting.

Relationship between predictors and the outcome variable

Before selecting the classification model, we assessed the strength of the relationship between the selected features and the outcome variable (HDP) using the Pearson correlation method.

Model training

The dataset was divided into a 90:10 ratio for training and testing purposes, ensuring sufficient data for model development while reserving a subset for evaluation. Five classification models were assessed for their performance: KNN, LR, SVM, RF, and XGBoost. Each model was trained on the training dataset, and its accuracy was calculated and compared to determine predictive performance. The classification model with the highest accuracy was subsequently selected and further validated for inclusion in the final predictive model development. In addition, we reported features importance on model prediction.

Model application

We downloaded new data from UCS Central Repository covering the period between April and May 2024. This new data has the same structure as previously downloaded data. Unlike the previous approach, we did not arithmetically calculate the risk of HDP using classical definitions but rather applied the developed model to predict the risk of HDP. We tested the prediction capabilities and reported the results.

Results

Descriptive results

Table 1. A total of 337,027 ANC visits were captured with visit dates ranging between the years 2020 and 2024. Over fifty percent of the visits appeared once and less than 5% of the visits appeared at least four times. The average gestational age during the visits was 24 weeks with a total of approximately eighty thousand visits (Figure 1). After we collapsed visits to

Table 1. Frequency distribution of ANC clients by visit/contact number before and after collapse.

	ANC visits (Before collapse)			ANC clients (After collapse)		
Visit No.	**Freq**	**Percent**	**Cum. Percent**	**Freq**	**Percent**	**Cum. Percent**
1	182,313	54.09%	54%	80,773	43.09%	43%
2	104,886	31.12%	85%	72,448	38.65%	82%
3	33,487	9.94%	95%	22,368	11.93%	94%
4	11,606	3.44%	99%	8,247	4.40%	98%
5	3,494	1.04%	100%	2,609	1.39%	99%
6	915	0.27%	100%	732	0.39%	100%
7	202	0.06%	100%	157	0.08%	100%
8	49	0.01%	100%	33	0.02%	100%
9	17	0.01%	100%	14	0.01%	100%
10	3	0.00%	100%	2	0.00%	100%
11	1	0.00%	100%	1	0.00%	100%
99	54	0.02%	100%	54	0.03%	100%
Total	**337,027**			**187,438**		

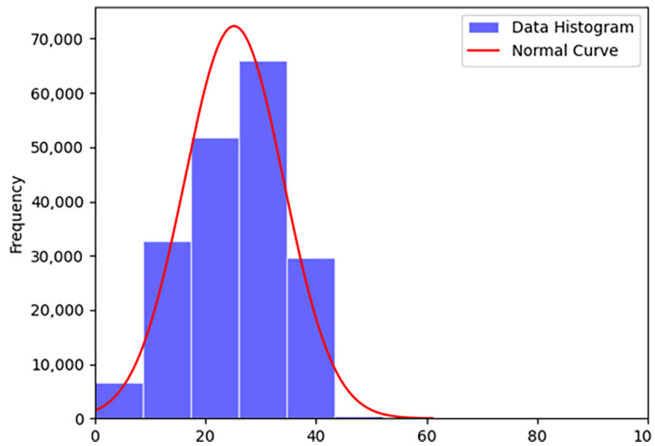

Figure 1. Frequency distribution by gestational age.

Table 2. Classification report.

	Precision	Recall	F1-score	Support
No Risk (0)	1	1	1	197
Risk (1)	1	1	1	182
Accuracy			1	379
Macro avg	1	1	1	379
Weighted avg	1	1	1	379

individual records, we obtained a total of 187,438 clients. Similar to visit details, the majority of the clients (43%) had a single visit, 6.3% of the clients had at least four visits (Table 2).

Correlation results

The results of the Pearson correlation are shown in Figure 2 below highlighting both systolic and diastolic pressure having significant positive correlation with HDP ($r = 0.69$, p-value < 0.05 and $r = 0.72$, $p < 0.05$ respectively).

Model training results

After data cleaning, we obtained a balanced dataset of 1,894 by 1,894 records with risk and no-risk of HDP. A 90-10% split gave us a total of 3,409 training records and 379 testing records. Cross-validation against five ML classification models (Figure 3) shows the following scores, KNN 97.27%, LG 93.52%, SVM 93.72%, RF 99.74% and XGB 99.88%. Based on the cross-validation scores, the XGB ML Classification class was selected as the strategy for the final model. Table 2 shows the performance of the SGB model using the test dataset.

Features importance

The results of the features importance is shown on Figure 4 with the base value of E [f(x)] = 0.152. Both Systolic BP (128mmHg) and Diastolic BP (67mmHg) values showed strong positive contribution to the final model output. The SHARP results were +7.94 and +7.62 respectively. The BMI and Protein in Urine showed small positive contributions (SHARP value of +0.89 and +0.24 respectively). Blood for Glucose (Value of 73) had a small negative contribution (SHARP value of -0.31). Temperature and Syphilis showed negligible impact to the overall prediction output.

Model application results

A total of 120,232 new data points were downloaded from the UCS central repository with similar data structure to the training dataset. When applied with the newly formulated model, the prediction results showed a total of 12,603 records with risk of HDP (107,629 records indicated no risk). Summary of the classification report is shown in Table 3, indicating an overall 90.95% model accuracy with 14% precision and 24% sensitivity (f1-score) for prediction Risk of HDP. The Area Under the Curve (AUC) score was 0.95 (95% CI: 0.9532 – 0.9548), indicating strong discriminatory power between the risk and no-risk group of HDP.

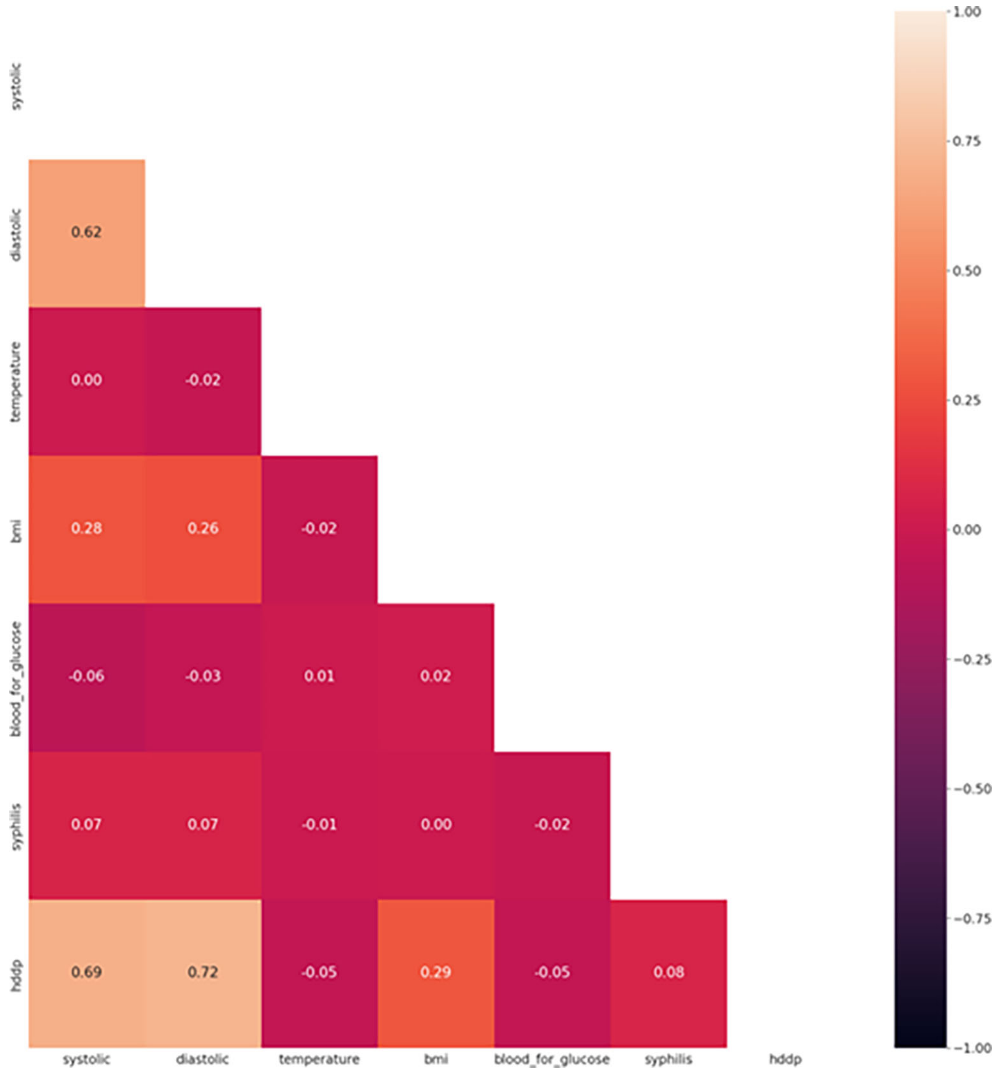

Figure 2. Heatmap showing correction of selected variables.

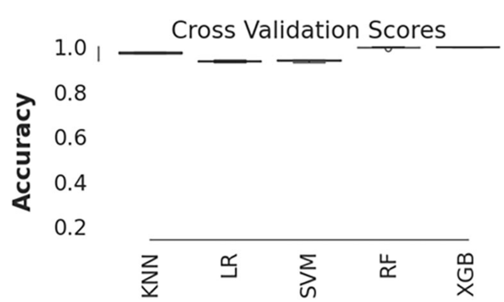

Figure 3. Cross validation results.

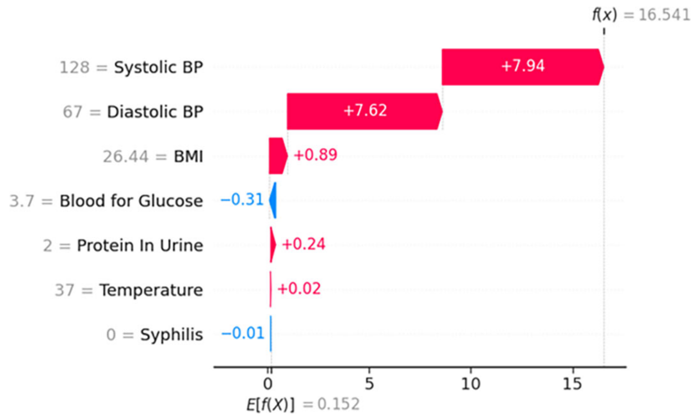

Figure 4. Features importance.

Table 3. Classification report.

	Precision	Recall	F1-score	Support
No Risk (0)	1	0.91	0.95	118,507
Risk (1)	0.14	1	0.24	1,725
Accuracy			0.91	120,232
Macro avg	0.57	0.95	0.6	120,232
Weighted avg	0.99	0.91	0.94	120,232

Discussion

ML models show significant potential for predicting adverse health outcomes and informing interventions, but their application in resource-limited settings remains constrained.[13–15] These challenges stem primarily from the lack of high-quality, readily available data, limited human resources to operationalize ML, and insufficient access to digital devices necessary for deployment.[16,17]

Despite these challenges, the increasing digital awareness, smartphone penetration, and government investment such as the UCS in Tanzania, are driving digital transformation in low-resource settings.[11] These advancements not only improve data availability, but also open new possibilities for integrating machine learning (ML) into routine healthcare services.[18] While routine data collection still requires quality improvements, this study demonstrates that existing data can effectively support ML models for maternal health risk prediction. By applying ML models, this study identified 10,878 cases at risk of hypertensive disorders in pregnancy (HDP)—cases that might have gone unnoticed in standard clinical routines. While not all these individuals may develop HDP, such predictive insights enable risk stratification, allowing healthcare providers to prioritize high-risk cases and implement targeted interventions as noted in other studies.[13,18–20] This proactive approach has the potential to improve maternal outcomes, optimize resource allocation, and strengthen the overall efficiency of healthcare delivery in low-resource settings.

In many low-resource settings, health data is often collected on paper, with poor quality and limited integration of data use at the source, thereby hindering validation and improvement processes.[21] Deploying digital solutions such as the UCS, processes and data collections can be streamlined to enhance data quality and facilitate data availability. However, the high cost and limited accessibility of such devices remain barriers. Moreover, human resources often lack adequate training and follow-up mechanisms, which impedes effective digital adoption.[17,22]

ML has immense potential as a decision-support tool for risk stratification, particularly in constrained environments where healthcare resources and workforce capacity are limited.[23,24] The model results indicate 91% accuracy but only 14% precision, reflecting a high number of false positives - commonly seen when aiming to minimize missed cases in high-stakes contexts.[25] However, even with low precision, the model's ability to flag high-risk cases can be valuable in prioritizing maternal care, optimizing resource allocation, and enabling timely interventions, ultimately improving maternal and newborn health outcomes in low-resource settings.[13]

This implementation highlights ML's ability to detect hidden patterns and relationships in clinical data, enhancing risk prediction and decision-making. While national guidelines for managing hypertensive HDP focus primarily on systolic and diastolic pressure, the ML model identified additional predictors - such as body mass index (BMI), protein in urine (positive association), and blood glucose levels (negative association)—which are routinely available but often overlooked in traditional assessments.[26] These subtle yet meaningful contributions may escape human observation but play a crucial role in refining risk stratification. This underscores ML's transformative potential as a decision-support tool, enabling more comprehensive, data-driven approaches to maternal health management. By integrating ML into routine care, healthcare providers can enhance early detection, optimize interventions, and ultimately improve maternal and newborn outcomes in resource-limited settings.

Limitations

A primary limitation of this study was the absence of a standardized method for recording HDP within the UCS. Although danger signs such as a history of hypertension, visual disturbances, severe headaches, and sudden weight gain were documented, they were inconsistently recorded and frequently omitted. This inconsistency underscores broader systemic challenges in routine ANC data collection, particularly in low-resource settings, where clinical documentation often lacks uniform formats and standardized terminologies.

Consequently, the training dataset did not include a clinically confirmed outcome variable for HDP. Instead, HDP cases were algorithmically defined using recorded blood pressure values, with hypertension classified as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg, in accordance with national ANC guidelines. While this approach aligns with accepted clinical definitions, the reliance solely on numeric blood pressure entries—without physician-confirmed diagnoses—may have affected the model's performance in terms of precision, sensitivity, and specificity.

Additionally, this study was retrospective in nature and utilized fully de-identified data extracted from routine ANC service records. Due to the absence of personally identifiable information, it was not possible to conduct follow-up assessments or verify outcomes through clinical record reviews. This limitation restricted our ability to evaluate the long-term predictive validity of the model or to confirm the accuracy of HDP classifications beyond the available dataset.

Recommendations

Governments in countries with low-resource settings need to create sustainable digital solutions incorporated with data use and data quality controls. Application of ML solutions should be considered, with efforts put in training and capacity building to ensure sustainable implementations.

Key messages

What is already known on this topic: Low-resource settings experience a high burden of maternal complications, including hypertensive disorders during pregnancy. Early detection is critical for timely interventions, improving maternal outcomes, and reducing mortality rates.

What this study adds: This study demonstrates the potential of machine learning (ML) models, integrated with digital health technologies, to predict the risk of hypertensive disorders in pregnancy using routine antenatal care (ANC) data. By leveraging ML, early identification of high-risk pregnancies is possible, even in settings with limited clinical resources.

How this study might affect research, practice, or policy: The findings highlight the need for policy reforms to enhance digital health infrastructure and standardize ANC data collection. Integrating ML-driven risk prediction tools into routine maternal healthcare could improve early detection, optimize resource allocation, and support data-driven decision-making in low-resource settings.

Author contributions

All authors made a significant contribution to this work. EM wrote the machine learning code and analysis. IL created the first draft. YK and AK reviewed and wrote the technical content. AH reviewed the draft, put references, and improved the discussion chapter. IL cleaned and created the final version. All authors took part in the interpretation and findings. RB, NK, PS, AM, JT, OS and AH took part to critically review the manuscript. All authors gave final approval of the version to be published.

Ethical consideration and consent

This study received ethical approval from the Institutional Review Board of the National Institute for Medical Research (NIMR), Tanzania, under approval number NIMR/HQ/R.8a/Vol.IX/4194 dated January 23rd, 2023. Informed verbal consent was obtained from all participants prior to conducting the interviews. Additional data were retrospectively sourced from records of routine antenatal care (ANC) clinic visits. All procedures were conducted in accordance with national ethical guidelines and regulations governing research involving human subjects.

Data availability statement

The data supporting the findings of this study are not publicly available due to institutional and national data protection policies. Access to the dataset may be granted upon reasonable request and with prior approval from the Ministry of Health, Tanzania. Interested researchers may contact the Ministry of Health to initiate the data access process, subject to applicable ethical and administrative approvals.

Acknowledgments

The authors would like to express their gratitude to the Ministry of Health, Tanzania, for granting access to and permission to use the data.

References

1. Montgomery-Csobán T, Kavanagh K, Murray P, *et al.*: **Machine learning-enabled maternal risk assessment for women with pre-eclampsia (the PIERS-ML model): a modelling study.** *The Lancet Digital Health.* 2024; **6**(4): e238–e250. PubMed Abstract | Publisher Full Text | Free Full Text
2. Li Y, Shen X, Yang C, *et al.*: **Novel electronic health records applied for prediction of pre-eclampsia: Machine-learning algorithms.** *Pregnancy Hypertens.* 2021; **26**: 102–109. PubMed Abstract | Publisher Full Text
3. Islam MN, Mustafina SN, Mahmud T, *et al.*: **Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda.** *BMC Pregnancy Childbirth.* 2022; **22**(1): 348. PubMed Abstract | Publisher Full Text | Free Full Text
4. Bertini A, Salas R, Chabert S, *et al.*: **Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review.** *Front. Bioeng. Biotechnol.* 2022 [cited 2023 Mar 24]; **9**. PubMed Abstract | Publisher Full Text | Free Full Text
5. Davidson L, Boland MR: **Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes.** *Brief. Bioinform.* [cited 2024 Nov 22]; **22**. Oxford Academic. PubMed Abstract | Publisher Full Text | Free Full Text
6. World Health Organization and others: *Evaluating the quality of care for severe pregnancy complications: the WHO near-miss approach for maternal health.* World Health Organization; 2011.
7. Berrueta M, Bardach A, Ciaponni A, *et al.*: **Maternal and neonatal data collection systems in low- and middle-income countries: scoping review protocol.** *Gates Open Res.* 2020; **4**: 18. PubMed Abstract | Publisher Full Text | Free Full Text
8. World Health Organization (WHO): *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience.* Geneva, Switzerland: WHO; 2018. Reference Source
9. Merid MW, Chilot D, Yigzaw ZA, *et al.*: **Women in low- and middle-income countries receive antenatal care at health institutions, yet not delivered there: a multilevel analysis of 2016–2021 DHS data.** *Trop. Med. Health.* 2024; **52**(1): 1. PubMed Abstract | Publisher Full Text | Free Full Text
10. **Association between achieving adequate antenatal care and health-seeking behaviors: A study of Demographic and Health Surveys in 47 low- and middle-income countries.** *PLoS Med.* [cited 2025 Feb 28]. Reference Source
11. *The United Republic of Tanzania Ministry of Health and Social Welfare. Tanzania National eHealth Strategy 2013 – 2018.* Ministry of Health and Social Welfare; [cited 2025 Feb 28]. Reference Source
12. **Unified Community System empowers health care workers and transforms communities in Tanzania.** *Digital Square.* 2023 [cited 2025 Mar 25]. Reference Source
13. Al Mashrafi SS, Tafakori L, Abdollahian M: **Predicting maternal risk level using machine learning models.** *BMC Pregnancy Childbirth.* 2024; **24**(1): 820. PubMed Abstract | Publisher Full Text | Free Full Text
14. Tomlinson M, Rotheram-Borus MJ, Swartz L, *et al.*: **Scaling up mHealth: where is the evidence?** *PLoS Med.* 2013; **10**(2): e1001382. PubMed Abstract | Publisher Full Text | Free Full Text
15. World Health Organization: *WHO guideline: recommendations on digital interventions for health system strengthening.* WHO; 2019. Reference Source
16. Mehl G, Labrique A: **Prioritizing integrated mHealth strategies for universal health coverage.** *Science.* 2014; **345**(6202): 1284–1287. PubMed Abstract | Publisher Full Text
17. Mumtaz H, Riaz MH, Wajid H, *et al.*: **Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review.** *Front. Digit. Health.* 2023 [cited 2025 Apr 5]; **5**. PubMed Abstract | Publisher Full Text | Free Full Text
18. Maleki Varnosfaderani S, Forouzanfar M: **The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century.** *Bioengineering.* 2024; **11**(4): 337. PubMed Abstract | Publisher Full Text | Free Full Text
19. Tomar K, Sharma CM, Prasad T, *et al.*: **A Machine Learning-Based Risk Prediction Model During Pregnancy in Low-Resource Settings.** *Med. Sci. Forum.* 2024; **25**(1): 13.
20. Yimer NB, Mekonnen EG: **Development and validation of a risk prediction model for caesarean delivery among multiparous women.** *Sci. Rep.* 2025; **15**(1): 5326. PubMed Abstract | Publisher Full Text | Free Full Text
21. Rose AL, Fenelon DL, Fils-Aimé JR, *et al.*: **Development of an Innovative Digital Data Collection System for Routine Mental Health Care Delivery in Rural Haiti.** *Glob. Health Sci. Pract.* 2021; **9**(4): 990–999.
22. Khan R, Khan S, Almohaimeed HM, *et al.*: **Utilization, challenges, and training needs of digital health technologies: Perspectives from healthcare professionals.** *Int. J. Med. Inform.* 2025; **197**: 105833. PubMed Abstract | Publisher Full Text
23. Ding L, Yin X, Wen G, *et al.*: **Prediction of preterm birth using machine learning: a comprehensive analysis based on large-scale preschool children survey data in Shenzhen of China.** *BMC Pregnancy Childbirth.* 2024; **24**(1): 810. PubMed Abstract | Publisher Full Text | Free Full Text
24. Habehh H, Gohel S: **Machine Learning in Healthcare.** *Curr. Genomics.* 2021; **22**(4): 291–300. PubMed Abstract | Publisher Full Text | Free Full Text
25. **Classification: Accuracy, recall, precision, and related metrics.** *Mach. Learn.* [cited 2025 Apr 5]. Google for Developers. Reference Source
26. **Gestational Hypertension and Preeclampsia: ACOG Practice Bulletin, Number 222.** *Obstet. Gynecol.* 2020; **135**(6): e237–e260. Publisher Full Text