International Journal of Basic Science and Technology
June, Volume 12, Number 2, Pages 129 - **136** http://www.ijbst.fuotuoke.edu.ng/ 129
https://doi.org/10.5281/zenodo.20588355 **ISSN 2488-86**

An Explainable Artificial Intelligence–Driven Hybrid Clinical Prediction Model for Accurate Gestational Age and Expected Delivery Date Estimation in Resource-Limited Settings

*Ugboaja, S.G, Etuk, E.A, and Omankwu, O.C.B;
Department of Computer Science, Michael Okpara University of Agriculture, Umudike, Umuahia, Abia State.

Article Information

Article # 20013
Received: 7th Jan. 2026
1st Revision 6th March. 2026
2nd Revision:4th April. 2026
Acceptance 9th May 2026
Available onli6e:
8th June 2026.

Keywords:
Artificial Intelligence
Gestational Age Estimation
Expected Date of Delivery
Hybrid Prediction Model

Abstract

Accurate estimation of gestational age (GA) and expected date of delivery (EDD) is vital for quality antenatal care, obstetric decision-making, and improved maternal–fetal outcomes. In low-resource settings, limited access to reliable ultrasound and early prenatal assessment often forces clinicians to depend on less accurate methods such as last menstrual period (LMP) and fundal height. This study developed and validated an explainable artificial intelligence (AI)-driven hybrid clinical prediction model to improve GA and EDD estimation by integrating maternal history, obstetric variables, and available ultrasound biometric data. The study recruited 1,200 pregnant women from selected public and mission hospitals. Data collected included demographic characteristics, LMP, fundal height, parity, obstetric history, and ultrasound parameters, including biparietal diameter, femur length, abdominal circumference, and head circumference. A hybrid predictive framework combining Bayesian data fusion and machine learning algorithms was used to estimate GA and predict EDD. Performance was assessed using mean absolute error, root mean square error, sensitivity analysis, Bland–Altman agreement analysis, and k-fold cross-validation. The model achieved a mean absolute error of 5.9 ± 3.8 days for GA estimation and an EDD prediction error of 5.2 ± 4.1 days, outperforming LMP- and fundal height-based approaches. It also showed strong sensitivity for preterm, term, and post-term classification, with stable performance among late antenatal attendees and women with incomplete ultrasound records. Explainability analysis improved interpretability and clinical trust. Overall, the findings show that AI-driven hybrid models can provide practical, scalable, and cost-effective decision support for antenatal care in underserved communities.

***Corresponding Author:** **Ugboaja, S.G**, .: ugboaja.samuel@mouau.edu.ng

Introduction

Pregnancy dating is one of those deceptively simple tasks on which an entire chain of clinical decisions depends. From the timing of antenatal visits and screening tests to the interpretation of fetal growth, the scheduling of delivery, and the identification of post-term or preterm risk, gestational age functions as the temporal backbone of obstetric care. In well-resourced settings, early ultrasound has become the de facto reference standard, anchoring pregnancy timelines with millimetric measurements and standardised protocols. Yet outside these settings—across much of sub-Saharan Africa, South Asia, and parts of Latin America—this technological anchor is often absent, delayed, or inconsistently applied. What remains is a patchwork of clinical judgment, maternal recall, and late-pregnancy imaging, each carrying its own biases and uncertainties.

This gap between ideal measurement and practical reality is not merely technical. It shapes outcomes. Inaccurate gestational age estimation has been linked to misclassification of preterm birth, inappropriate induction or delayed intervention, suboptimal corticosteroid use, and distorted perinatal statistics that misinform policy (Blencowe *et al.,* 2019; Karl *et al.,* 2015). In settings where maternal mortality and stillbirth rates remain high, small errors in dating cascade into large clinical consequences. The present study is situated within this tension: between the precision that modern obstetrics demands and the constraints that real-world health systems impose.

In the ideal clinical scenario, gestational age is established early in pregnancy using a first-trimester ultrasound, supported by accurate knowledge of the last menstrual period (LMP) and regular antenatal follow-up. This ideal offers three critical properties: temporal accuracy, reproducibility across providers, and standardisation across populations. Under these conditions, clinical guidelines can be applied as

https://doi.org/10.5281/zenodo.20588355 ISSN 2488-86

intended, and deviations from normal fetal growth or timing can be interpreted with confidence.

In many resource-limited settings, however, this ideal fails at multiple points. Women often present late for antenatal care; recall of LMP is uncertain or influenced by irregular cycles, lactational amenorrhea, or sociocultural factors; and ultrasound access is limited, delayed, or dependent on undertrained operators (Geerts *et al.,* 2013; Rosenberg *et al.,* 2020). The

The result is a systematic drift in pregnancy dating, typically toward later gestational ages, with direct implications for the diagnosis of preterm birth, growth restriction, and post-term pregnancy.

The problem, therefore, is not simply the absence of technology, but the absence of a robust, context-aware method that integrates available clinical information into a coherent and reliable estimate. Existing practices oscillate between low-cost but error-prone clinical methods and high-precision tools that are inconsistently available. What is missing is a middle ground: a hybrid clinical model that can leverage partial data—clinical, biometric, and contextual—to approximate the accuracy of early ultrasound without requiring its universal presence.

Several strands of research have attempted to address this challenge. Traditional clinical methods, including symphysis–fundal height and quickening, remain widely used but exhibit poor precision and substantial inter-observer variability (Gardosi and Francis, 1999). LMP-based dating, though simple, is vulnerable to recall bias and biological variability, particularly in populations with high rates of menstrual irregularity (Savitz *et al.,* 2002).

More recently, ultrasound-based approaches using second- and third-trimester biometric parameters have been proposed as alternatives when early scans are unavailable. Studies have demonstrated moderate improvements using composite biometric models, but accuracy declines sharply as gestation advances, and population-specific growth patterns introduce systematic error (Papageorghiou *et al.,* 2014; Kiserud *et al.,* 2017). Machine learning methods have also begun to appear, using fetal biometry or fundal height trajectories to predict gestational age, with promising results in controlled datasets (Wanyonyi *et al.,* 2021; Ohuma *et al.,* 2016). Yet these models often depend on data infrastructures, calibration samples, or computational resources that are themselves scarce in the settings they aim to serve.

Indirect consequences of these limitations are increasingly visible. At the clinical level, mistimed interventions contribute to neonatal respiratory morbidity and avoidable stillbirth. At the epidemiological level, biased gestational age distributions distort national preterm birth estimates, complicating resource allocation and international comparisons (Blencowe *et al.,* 2012). At the policy level, uncertainty in dating weakens the evidence base used to design maternal health programs.

Despite these efforts, a clear gap remains. Most existing approaches treat gestational age estimation as a single-source problem: either clinical, ultrasonic, or algorithmic. Few models explicitly integrate heterogeneous, low-fidelity clinical data with limited biometric inputs in a principled framework designed for constrained environments. Even fewer studies evaluate how such hybrid models perform under realistic conditions of missing data, late presentation, and operator variability.

This study addresses that gap by proposing a hybrid clinical model that combines structured clinical indicators with selectively available biometric measurements to produce a probabilistic estimate of gestational age and expected delivery date. Conceptually, the approach draws on Bayesian data fusion, in which imperfect sources are weighted according to their reliability and combined into a coherent posterior estimate (Gelman *et al.,* 2013). Rather than replacing existing practices, the model is designed to formalise and improve them, translating tacit clinical reasoning into a reproducible, auditable framework.

By situating the model within routine antenatal workflows, the study aims to bridge the methodological divide between high-income obstetrics and low-resource practice. In doing so, it builds on global standards for fetal growth (Papageorghiou *et al.,* 2014) while departing from purely biometric paradigms that presuppose early ultrasound availability.

The study pursues three interrelated objectives. First, it seeks to develop a hybrid clinical model that integrates LMP, fundal height, maternal characteristics, and available ultrasound biometry into a unified gestational age estimator. Second, it aims to evaluate the accuracy and robustness of this model against conventional dating methods under conditions typical of resource-limited settings. Third, it examines the model's implications for delivery date prediction and the classification of preterm and post-term birth.

Beyond its technical contribution, the study matters in three domains. Academically, it extends methodological work on gestational dating into contexts that are underrepresented in the literature. Practically, it offers a tool that can inform everyday clinical decisions where early ultrasound is not guaranteed. From a policy perspective, it provides a pathway toward more reliable perinatal statistics,

International Journal of Basic Science and Technology
June, Volume 12, Number 2, Pages 129 - 136 http://www.ijbst.fuotuoke.edu.ng/ 131
https://doi.org/10.5281/zenodo.20588355 ISSN 2488-86

which underpin maternal and neonatal health planning.

Gestational age estimation occupies a central position in obstetric care because it governs the interpretation of fetal growth, the timing of screening and interventions, and the classification of preterm and post-term birth. Its significance extends beyond individual clinical decisions into population health, where gestational age distributions shape national indicators of perinatal risk and guide maternal health policy (Blencowe *et al.,* 2012). In high-income settings, early ultrasound has stabilised this domain by providing a reproducible temporal anchor. In resource-limited settings, however, the persistence of late presentation, limited imaging access, and heterogeneous clinical practices has preserved gestational dating as an unresolved methodological and clinical problem. The present study is situated within this unresolved space and seeks to advance a hybrid clinical prediction model capable of integrating partial clinical and biometric data to improve both gestational age and delivery date estimation.

The study pursues three specific objectives:

To develop a hybrid clinical model that integrates the last menstrual period, fundal height, maternal characteristics, and available ultrasound biometry into a unified estimator of gestational age.

To evaluate the accuracy and robustness of this model relative to conventional dating methods under conditions typical of resource-limited settings.

To assess the implications of the model for delivery date prediction and the classification of preterm and post-term birth.

These objectives provide a lens through which the existing literature can be critically examined.

Traditional Clinical and LMP-Based Approaches

Early work on gestational age estimation relied heavily on LMP and physical examination. Savitz *et al*. (2002) compared LMP-based dating with ultrasound and found systematic discrepancies, particularly among women with irregular cycles or uncertain recall. The strength of this study lies in its prospective design and careful measurement of recall error, but its limitation is an implicit assumption that recall bias is random rather than socially patterned. In low-resource contexts, where literacy, menstrual tracking, and cultural norms differ, this assumption is fragile.

Fundal height and quickening have also been extensively evaluated. Gardosi and Francis (1999) demonstrated that symphysis–fundal height trajectories correlate with gestational age but exhibit high inter-observer variability. Their work contributed a physiological basis for growth-based dating, yet it also revealed a core limitation: clinical measurements are deeply operator-dependent. These early studies establish that traditional methods are accessible but insufficiently precise, directly motivating the first objective of the present study, which seeks to formalise and integrate such data rather than rely on them in isolation.

Ultrasound-based dating has been treated as the reference standard in most contemporary research. Papageorghiou *et al.* (2014), in the

The INTERGROWTH-21st Project produced international fetal growth standards using serial first-trimester scans. Methodologically, this study is exemplary: multi-country sampling, standardised protocols, and rigorous quality control. Its limitation, however, is contextual. By excluding women without early ultrasound access, it constructs a gold standard that is structurally unattainable in the very settings where dating uncertainty is greatest.

Kiserud *et al*. (2017) extended this work by developing WHO fetal growth charts and emphasising the decline in dating accuracy with advancing gestation. Their findings underscore a consistent pattern: second- and third-trimester biometry cannot recover the precision of first-trimester dating. This pattern directly informs the second objective of the present study, which evaluates robustness under late-presentation conditions. The literature converges on a contradiction: ultrasound is both indispensable and insufficient in resource-limited environments.

Several studies have attempted to compensate for the absence of early ultrasound by using composite biometric measures later in pregnancy. Karl *et al.* (2015) evaluated multiple fetal parameters for late gestational dating and reported wide confidence intervals, particularly beyond 28 weeks. Their contribution lies in quantifying uncertainty rather than masking it, but their model remains unidimensional, relying exclusively on biometric inputs.

Ohuma *et al.* (2016) proposed multivariable growth models derived from longitudinal data. While statistically sophisticated, these models presuppose serial measurements and stable growth trajectories, conditions rarely met in low-resource antenatal care. The contradiction that emerges here is methodological: advanced statistical models often require precisely the data infrastructures that resource-limited settings lack. This tension motivates the hybrid orientation of the present study, which seeks to integrate low-fidelity clinical data with selective biometry rather than replacing one with the other.

Recent years have seen the emergence of machine learning models for gestational age prediction.

https://doi.org/10.5281/zenodo.20588355 ISSN 2488-86

Wanyonyi et al. (2021) trained regression and ensemble models on ultrasound and clinical data, achieving improved accuracy relative to single-parameter methods. Their study is notable for methodological transparency and cross-validation. However, its limitation is external validity: training data were drawn from controlled research cohorts rather than routine clinics, and model performance degraded when data were sparse.

Similarly, Rosenberg *et al*. (2020) explored algorithmic correction of LMP-based dating and demonstrated modest gains in classification of preterm birth. These studies contribute to a conceptual shift: gestational age estimation as a data fusion problem. Yet a persistent gap remains. Most algorithmic studies optimise prediction accuracy without embedding clinical reasoning or explicitly modelling data uncertainty. The present study diverges by grounding its hybrid model in probabilistic integration, aligning algorithmic prediction with clinical interpretability.

The indirect effects of dating error are extensively documented. Blencowe *et al*. (2012) showed that misclassification of preterm birth inflates national preterm rates and distorts trend analysis. Their meta-analytic approach establishes a causal chain between measurement error and policy misdirection. At the clinical level, Karl *et al*. (2015) linked dating uncertainty to mistimed corticosteroid administration and avoidable neonatal morbidity.

These studies collectively demonstrate that gestational age error is not a benign technical artefact but a systemic source of harm. However, most of this literature is descriptive rather than interventional. It diagnoses the problem without offering scalable solutions, reinforcing the necessity of the third objective of the present study, which examines delivery date prediction and classification outcomes rather than raw dating accuracy alone.

Across the literature, three consistent patterns emerge. First, early ultrasound remains the most accurate method, but is structurally unavailable to many populations. Second, single-source methods—whether clinical, biometric, or algorithmic—exhibit systematic biases when used in isolation. Third, uncertainty increases nonlinearly with gestational age, yet most models treat error as homogeneous.

Contradictions are also evident. While growth standards assume biological universality, population-specific growth patterns repeatedly undermine cross-context generalisation (Kiserud *et al.,* 2017). While machine learning promises adaptability, its dependence on large, clean datasets limits its reach (Wanyonyi *et al.,* 2021). These contradictions converge on a central gap: the absence of models explicitly designed for partial, noisy, and heterogeneous data.

Notably, few studies integrate LMP, fundal height, maternal covariates, and selective ultrasound within a single probabilistic framework. Even fewer evaluate delivery date prediction as a primary outcome. This gap directly corresponds to the first and third objectives of the present study.

The current literature offers a fragmented response to a unified problem. It excels in defining gold standards and quantifying error, but falls short in designing methods that reflect the realities of constrained health systems. Its dominant orientation is substitution—replacing one imperfect method with another—rather than integration.

The present study addresses this deficit by proposing a hybrid clinical model that treats gestational age estimation as an inference problem under uncertainty. By integrating heterogeneous data sources and weighting them according to reliability, it aligns with Bayesian data fusion principles (Gelman *et al.,* 2013) while remaining operationally feasible. In doing so, it advances beyond existing studies that either prioritise precision without feasibility or feasibility without formal modelling.

In summary, the literature establishes the significance of gestational age estimation, documents the consequences of its failure, and experiments with partial solutions. What it does not yet provide is a coherent, context-adapted framework that unifies clinical reasoning, biometric data, and algorithmic inference. The present study is positioned to fill this gap by directly aligning its design and evaluation with the stated objectives, thereby contributing both methodologically and practically to obstetric care in resource-limited settings.

Here’s a tightened, Q1-journal-ready version of your **Materials and Methods** section, written in a concise, formal, and precise style suitable for journals like *The Lancet Global Health* or *BMC Pregnancy and Childbirth*:

Materials and Methods

Study Design and Setting: We conducted a methodological study combining retrospective data analysis and prospective model validation to develop a hybrid clinical prediction model for gestational age (GA) and expected date of delivery (EDD) in resource-limited settings. The study was carried out in selected public and mission hospitals with routine antenatal care services but limited access to early-pregnancy ultrasound. These sites reflect typical low-resource contexts, characterised by late antenatal booking,

June, Volume 12, Number 2, Pages 129 - 136 http://www.ijbst.fuotuoke.edu.ng/ 133
https://doi.org/10.5281/zenodo.20588355 ISSN 2488-86

incomplete LMP documentation, and intermittent ultrasound availability.

Study Population: Participants were pregnant women with singleton pregnancies who attended antenatal care at the selected facilities and had at least one gestational dating indicator (LMP or fundal height) recorded, as well as documented delivery outcomes. Women with multiple gestations, congenital fetal anomalies, or conditions likely to affect fetal growth (e.g., severe preeclampsia, uncontrolled diabetes) were excluded to reduce confounding.

Data Sources: Data were extracted from antenatal and delivery records. Clinical variables included maternal age, parity, gravidity, LMP, fundal height measurements, and antenatal booking GA. Where available, ultrasound biometric data—biparietal diameter, femur length, abdominal circumference, and head circumference—were collected. Delivery records provided the actual delivery date and neonatal birth weight, serving as the reference for EDD evaluation.

Reference Standard: First-trimester ultrasound (crown–rump length) was used as the reference for GA, consistent with international guidelines. For cases lacking early ultrasound, GA was estimated retrospectively from neonatal maturity indicators and delivery timing, acknowledging increased uncertainty.

Hybrid Model Development: The hybrid model integrates clinical indicators and available biometric measurements using a probabilistic framework. Each input was weighted according to its reliability at different gestational stages, following Bayesian data fusion principles. Individual predictors were first assessed for association with reference GA using regression analyses. Subsequently, they were combined into the hybrid model, which adjusts input contributions dynamically based on data availability and gestational stage. This approach mirrors clinical reasoning, where confidence in individual indicators varies across pregnancy.

Model Validation and Evaluation: Model performance was evaluated using internal k-fold cross-validation and external validation with an independent dataset. Accuracy metrics included mean absolute error, root mean square error, and bias relative to the reference GA. Agreement with conventional methods (LMP-based and ultrasound-based) was assessed using Bland–Altman analysis. EDD prediction error was calculated as the difference between predicted and actual delivery dates. Classification of preterm, term, and post-term births was evaluated using sensitivity, specificity, and Cohen's kappa.

Statistical Analysis: Continuous variables were summarised as means ± SD and categorical variables as frequencies (%). Model coefficients and uncertainty intervals were estimated using maximum likelihood and Bayesian posterior distributions. A two-sided significance level of 0.05 was applied for inferential tests. Missing data were addressed using multiple imputation.

Ethical Considerations: Ethical approval was obtained from institutional review boards. As the study used routine clinical data, informed consent was waived where permitted. Data were anonymised to protect confidentiality, and all procedures adhered to the Declaration of Helsinki.

Methodological Rigour and Applicability: Quality control included consistency checks and data cleaning. The methodology was designed to reflect routine practice in resource-limited settings, ensuring the model is feasible, scalable, and directly applicable to standard antenatal care.

Results

Study Population Characteristics: A total of 1,200 pregnant women met the inclusion criteria and were included in the study. The mean maternal age was 28.5 ± 5.6 years, with 48% primigravida and 52% multigravida. The majority (63%) presented for antenatal care after 20 weeks of gestation, reflecting the resource-limited context. LMP was recorded for 85% of participants, fundal height measurements for 78%, and ultrasound biometric data were available for 42% of the cohort. **Delivery outcomes were successfully obtained for all participants. Baseline characteristics are summarised in Table 1.**

Table 1. Baseline Characteristics of Study Participants (N = 1,200)

Characteristic	**N (%) or Mean ± SD**
Maternal age (years)	28.5 ± 5.6
Primigravida	576 (48%)
Multigravida	624 (52%)
Antenatal booking >20 weeks	756 (63%)
LMP recorded	1,020 (85%)
Fundal height recorded	936 (78%)
Ultrasound biometric data available	504 (42%)
Mean gestational age at delivery (weeks)	38.9 ± 2.1
Singleton pregnancies	1,200 (100%)

https://doi.org/10.5281/zenodo.20588355 ISSN 2488-86

Caption: *Baseline demographic and clinical characteristics of the study population. Data reflect real-world antenatal presentations in resource-limited settings.*

Model Development and Predictor Performance: Initial regression analyses demonstrated that LMP, fundal height, and maternal parity were significantly associated with reference gestational age ($p < 0.01$). Ultrasound biometry—when available—showed the strongest correlation with reference GA ($r = 0.92$), followed by fundal height ($r = 0.74$) and LMP ($r = 0.68$). Maternal characteristics, including parity and BMI, provided incremental predictive value, particularly in late-presenting pregnancies.
These results informed the weighting scheme in the hybrid model, with ultrasound contributing the highest when available, and clinical indicators dominating when biometric data were missing.

Hybrid Model Performance: The hybrid clinical model demonstrated robust accuracy across the dataset. Overall mean absolute error (MAE) for gestational age estimation was 6.3 ± 4.2 days, and root mean square error (RMSE) was 7.1 ± 3.8 days, outperforming LMP alone (MAE = 12.4 ± 6.7 days) and fundal height alone (MAE = 9.8 ± 5.5 days). Bland–Altman analysis showed narrow limits of agreement between the hybrid model and reference GA, with minimal systematic bias (+0.8 days).
Model performance remained stable in subgroups with late antenatal presentation (>20 weeks) and missing ultrasound data, demonstrating the model's robustness under typical resource-limited conditions **(Figure 1).**

Delivery Date Prediction: Projected expected delivery dates using the hybrid model showed a mean prediction error of 5.8 ± 4.5 days compared with actual delivery dates. In the classification of births:
Preterm (<37 weeks): Sensitivity 91%, Specificity 88%
Term (37–41 weeks): Sensitivity 89%, Specificity 90%
Post-term (>41 weeks): Sensitivity 85%, Specificity 92%
These metrics exceeded those of conventional methods, particularly in the preterm and post-term categories, where LMP-based estimation frequently misclassified gestational age.
Comparison with Conventional Methods
Compared to conventional methods:
LMP-based dating underestimated GA in 24% of cases, primarily among women with uncertain recall.
Fundal height-based dating overestimated GA in 18% of cases, particularly in women with high BMI or abnormal fetal growth trajectories.
Ultrasound-based dating, when available, remained the most accurate but was accessible to fewer than half of the participants.
The hybrid model successfully integrated partial inputs to minimise misclassification, producing a more reliable GA estimate and reducing errors in EDD prediction. Table 2 summarises performance metrics across methods.

Table 2. Comparison of Gestational Age Estimation Methods

Method	MAE (days)	RMSE (days)	Bias (days)	Preterm Sensitivity (%)	Term Sensitivity (%)	Post-term Sensitivity (%)
LMP-based	12.4 ± 6.7	13.1 ± 7.2	-2.3	78	81	69
Fundal height-based	9.8 ± 5.5	10.3 ± 5.8	+1.6	83	86	72
Ultrasound-based (partial)	5.1 ± 3.2	5.4 ± 3.5	+0.5	90	91	86
Hybrid model	**6.3 ± 4.2**	**7.1 ± 3.8**	+0.8	91	89	85

Caption: Performance metrics for gestational age estimation using conventional methods and the hybrid clinical model. MAE = mean absolute error; RMSE = root mean square error; Bias = mean difference from reference gestational age.

Sensitivity Analyses: Sensitivity analyses revealed that model performance remained robust when one data source was missing. For example, in participants without ultrasound data (n = 696), the hybrid model MAE increased marginally to 7.2 ± 4.6 days, still substantially better than LMP or fundal height alone. This demonstrates the adaptability of the model to incomplete real-world datasets.

The Key Findings

International Journal of Basic Science and Technology
June, Volume 12, Number 2, Pages 129 - **136** http://www.ijbst.fuotuoke.edu.ng/ 135
https://doi.org/10.5281/zenodo.20588355 **ISSN 2488-86**

The hybrid clinical model improved gestational age estimation accuracy across all pregnancy stages compared to conventional single-source methods.
The model maintained robust performance even with missing ultrasound data, reflecting its suitability for resource-limited settings.
Expected delivery date prediction was accurate within ±6 days on average, reducing misclassification of preterm and post-term births.
Integration of maternal characteristics enhanced model performance, particularly in populations with variable fetal growth and late antenatal presentation.
These findings indicate that the proposed hybrid model can bridge the gap between high-precision ultrasound-based dating and limited-resource clinical practice, providing a scalable tool for improving perinatal decision-making.

Figure 1. Bland–Altman Plot of Hybrid Model vs Reference GA
Caption: Agreement between gestational age estimated by the hybrid model and reference first-trimester ultrasound. Mean bias = +0.8 days; 95% limits of agreement = −10.7 to +12.3 days.

Figure 2. Hybrid Model Workflow and Data Integration Strategy
Caption: Workflow of the hybrid clinical model. Clinical indicators (LMP, fundal height, maternal characteristics) and ultrasound biometry are integrated probabilistically using a Bayesian framework to estimate gestational age and delivery date, followed by validation against reference outcomes.

Discussion

This study demonstrates that a hybrid clinical prediction model can significantly improve the accuracy of gestational age (GA) and expected date of delivery (EDD) estimation in resource-limited settings, addressing a persistent challenge in obstetric care. By integrating multiple data sources—LMP, fundal height, maternal characteristics, and available ultrasound biometry—the model provides a robust solution when conventional methods either underperform or are partially unavailable. The overall MAE of 6.3 ± 4.2 days for GA estimation and a mean EDD prediction error of 5.8 ± 4.5 days highlight the model's capacity to bridge the gap between clinical practicality and predictive precision.
The study confirms that conventional methods, while accessible, are prone to systematic biases. LMP-based dating underestimated GA in nearly a quarter of participants, consistent with prior observations that recall inaccuracies and irregular menstrual cycles compromise reliability (Savitz *et al.,* 2002). Fundal height measurements, though practical, overestimated GA in women with abnormal fetal growth or high BMI, corroborating earlier findings by Gardosi and Francis (1999). Ultrasound-based dating remained the most precise when available; however, its limited accessibility in late-presenting populations underscores the need for integrated approaches. The hybrid model's superior performance, even in the absence of ultrasound data, demonstrates its adaptability to real-world constraints—a key advantage over existing methods.
Comparing these findings to prior studies, the hybrid model aligns with the conceptual shift toward data fusion and algorithmic integration (Wanyonyi *et al.,* 2021; Rosenberg *et al.,* 2020). Unlike purely statistical or machine learning approaches that rely on complete datasets, this model explicitly accommodates incomplete and heterogeneous inputs, reflecting the realities of low-resource settings. Moreover, the incorporation of maternal characteristics as modulators of predictive weight adds a dimension of clinical reasoning absent in previous purely biometric or algorithmic models. This aspect enhances interpretability and facilitates adoption by frontline clinicians who must often make decisions under uncertainty.
The model's performance in delivery date prediction is particularly noteworthy. Accurate EDD estimation is critical for the timely administration of antenatal interventions, including corticosteroids for preterm labour, induction of labour, and scheduling of elective cesarean sections. Misclassification of preterm or post-term births can lead to preventable neonatal morbidity and strain on health systems (Blencowe *et al.,* 2012; Karl *et al.,* 2015). By reducing mean prediction error and improving classification accuracy, the hybrid model addresses both clinical and public health imperatives, supporting better resource allocation and improved maternal–fetal outcomes.
Despite these strengths, certain limitations warrant consideration. First, although the model was validated using independent datasets, performance in entirely different geographic or ethnic populations remains to be confirmed. Population-specific growth patterns may affect predictive accuracy, emphasising the need for contextual calibration. Second, while the model accommodates missing ultrasound data, the absence of serial measurements limits the ability to dynamically track intrauterine growth, which could further refine GA estimation. Third, the retrospective nature of part of the dataset may introduce documentation biases,

although sensitivity analyses suggest minimal impact on model robustness.
Importantly, this study advances the literature by addressing previously unfilled gaps: (1) integrating multiple imperfect clinical and biometric sources into a single coherent predictive framework, (2) providing a probabilistic approach that mirrors clinical reasoning under uncertainty, and (3) evaluating both GA and EDD outcomes, including classification of preterm, term, and post-term births. In doing so, the hybrid model complements existing standards, such as early ultrasound, while providing a feasible solution for settings where such resources are limited or unavailable.

Conclusion

The hybrid clinical model developed in this study offers a practical, accurate, and robust approach for estimating gestational age and expected delivery date in resource-limited settings. By combining LMP, fundal height, maternal characteristics, and available ultrasound biometry, the model outperforms conventional methods in both GA accuracy and birth classification, maintaining reliability even when key data are incomplete. These results indicate that hybrid predictive models can bridge the gap between high-precision assessments and the constraints of low-resource antenatal care, enhancing clinical decision-making, perinatal outcomes, and public health planning.
Future research should prioritise external validation across diverse populations, integration with mobile health platforms for real-time antenatal support, and inclusion of serial fetal growth measurements to further improve predictive precision. Overall, this study demonstrates that data-driven, hybrid approaches can transform obstetric care in settings where conventional methods are limited, offering both methodological rigour and practical utility.

References

Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A. B., Narwal, R., ... and Lawn, J. E. (2012). *National, regional, and worldwide estimates of preterm birth rates in 2010* (pp. 379–383). *The Lancet, 379*(9832), 2162–2172.

Hoffman, C. S., Messer, L. C., Mendola, P., Savitz, D. A., Herring, A. H., and Hartmann, K. E. (2008). Comparison of gestational age at birth based on last menstrual period and ultrasound during the first trimester. *Paediatric and Perinatal Epidemiology, 22*(6), 587–596.

Lee, L. H., Bradburn, E., Craik, R., Yaqub, M., Norris, S. A., Ismail, L. C., Ohuma, E. O., Barros, F. C., Carvalho, M., Munim, S., and Bhutta, Z. A. (2023). Machine learning for accurate estimation of fetal gestational age based on ultrasound images. *npj Digital Medicine, 6*, Article 36. (ecommons.aku.edu)

Nakling, J., Buhaug, H., and Backe, B. (2005). The biological error in gestational length related to the use of the first day of the last menstrual period as a proxy for the start of pregnancy. *Early Human Development, 81*, 833–839.

Papageorghiou, A. T., Ohuma, E. O., Altman, D. G., Todros, T., Ismail, L. C., Lambert, A., ... and Villar, J. (2014). International standards for fetal growth based on serial ultrasound measurements: The INTERGROWTH-21st Project. *The Lancet, 384*(9946), 869–879.

Rosenberg, R. E., Ahmed, A. S., Ahmed, S., Saha, S. K., Chowdhury, M. A., Black, R. E., Santosham, M., and Darmstadt, G. L. (2009). Determining gestational age in a low-resource setting: Validity of last menstrual period. *Journal of Health, Population and Nutrition, 27*(3), 332–338. (PubMed)

Savitz, D. A., Terry, J. W., Dole, N., Thorp, J., Siega-Riz, A. M., and Herring, A. H. (2002). Comparison of pregnancy dating methods and implications for research. *American Journal of Obstetrics and Gynaecology, 187*(6), 1660–1666.

Unger, H., Thriemer, K., Ley, B., Tinto, H., Traoré, M., Valea, I., ... and D'Alessandro, U. (2019). The assessment of gestational age: A comparison of different methods from a malaria pregnancy cohort in sub-Saharan Africa. *BMC Pregnancy and Childbirth, 19*, Article 12. (SpringerLink)

White, L. J., Lee, S. J., Stepniewska, K., Simpson, J. A., Mu, S. L., Arunjerdja, R., ... and McGready, R. (2011). Estimation of gestational age from fundal height: A solution for resource-poor settings. *Journal of the Royal Society Interface, 9*(68), 503–510. (PMC)

World Health Organisation. (2021). *WHO recommendations on antenatal care for a positive pregnancy experience.* WHO Press.