JIRBDAI

Journal of Institutional Research, Big Data Analytics and Innovation

Predicting Unmet Contraceptive Need Among Adolescent Girls and Young Women in North-Central Nigeria: Using a Spatial Machine Learning Model

[1]Muhammad Farida Ghazali, [2]K. H. Oduwole, [3]M. O. Adenomon, [4]Kefas Yunana

Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria

Corresponding Author: ferhweederhghazali@gmail.com

ABSTRACT

Unmet need for contraception remains high among adolescent girls and young women in Nigeria, yet most Nigerian studies rely on conventional regression applied to older survey rounds and seldom ask where the burden falls. This study predicted unmet contraceptive need among adolescent girls and young women aged 15 to 24 years in North-Central Nigeria by combining spatial statistics with ensemble machine learning. Data came from the 2023 to 2024 Nigeria Demographic and Health Survey linked to its geographic cluster dataset, giving 1,290 young women exposed to the risk of pregnancy in 245 clusters. Spatial clustering was tested with Moran's I, local indicators of spatial association and the Getis-Ord Gi* statistic, with sensitivity analyses for coordinate displacement. A logistic regression baseline and seven tree-based and ensemble learners were compared by repeated cluster-grouped cross-validation and on a held-out test set, with outcome-based features built from training data only. Weighted unmet need was 22.1% (interval 19.3% to 25.0%), ranging from 17.4% in Niger to 35.8% in Kwara, where it was driven mainly by sexually active young women who had never married. Unmet need was spatially clustered (Moran's I = 0.0797, p = 0.013), with 15 significant hotspots, 13 of them in Kwara, a concentration that persisted across neighbourhood definitions and repeated displacement of the coordinates. The machine learning models reached a mean ROC-AUC of 0.684 to 0.713 against 0.637 for the baseline; the voting ensemble was highest, with a test ROC-AUC of 0.698, and the stacking ensemble gave the best-calibrated probabilities (calibration slope 1.03). Marital or union status and children ever born were the leading predictors. Spatial features built without outcome leakage added little to prediction (change in ROC-AUC 0.000 to +0.008). Unmet need is geographically concentrated and ensemble learning predicts it better than regression, but location adds little beyond individual and community characteristics. Programmes should prioritise the high-need area of southern Kwara, verified locally, and services for unmarried sexually active young women.

KEYWORDS: Contraception; Adolescent Girls, Young Women; Spatial Analysis; Ensemble Machine Learning; Hotspot Analysis; Nigeria

I. INTRODUCTION

Unmet need for contraception describes a woman who is fecund and sexually active, who wants to delay or stop childbearing, and who is not using any method (Bradley et al., 2012). It is the indicator that links reproductive intention to the services meant to support it, and in Nigeria the gap remains wide. The 2023 to 2024 Nigeria Demographic and Health Survey (NDHS) reports low contraceptive uptake alongside early childbearing (NPC et al., 2025), and a multilevel analysis of Nigerian adolescent girls and young women (AGYW) aged 15 to 24 years found modern

https://doi.org/10.5281/zenodo.24199907

contraceptive prevalence of only 9.1% (Muhammad et al., 2026). Young women carry the heaviest consequences of unintended pregnancy, from interrupted schooling to obstetric risk, which is why family planning programmes treat them as a priority segment (Solanke et al., 2022). The burden is not spread evenly. A spatio-temporal analysis of five Nigerian surveys from 2003 to 2024 found that contraceptive use rose only from 14.5% to 20.7% and that the same areas stayed disadvantaged throughout (Oladimeji et al., 2026). Among youth, modern contraceptive use was 12.6% in the south against 3.9% in the north (Sanni et al., 2025). If need clusters in particular places, programmes that treat a zone as uniform will spend resources where they are least needed. Planners therefore need two things at once: a map of where unmet need concentrates, and a model that identifies which young women are most likely to be affected.

North-Central Nigeria is a useful setting for producing both. The zone, made up of Benue, Kogi, Kwara, Nasarawa, Niger and Plateau states and the Federal Capital Territory (FCT), lies between the largely Muslim north and the largely Christian south, and it contains the highly urban FCT alongside deeply rural districts. This study combined formal spatial statistics with ensemble machine learning to predict unmet need among AGYW in the zone. It had four objectives: to test whether unmet need is spatially clustered; to identify the factors associated with it; to compare ensemble learners with a conventional logistic baseline; and to measure how much spatial information adds to prediction. Three gaps in the research stand in the way of better targeting of family planning for young women in Nigeria. First, Nigerian studies of unmet need rely almost entirely on conventional and multilevel logistic regression applied to the 2018 survey or earlier rounds (Oyinlola et al., 2024; Odjesa & Okonofua, 2024; Solanke et al., 2022). These models estimate average associations but are not built to predict which individuals are at risk. Second, the growing machine learning literature on African reproductive health has examined other outcomes and has rarely used the survey's geographic coordinates (Melaku et al., 2025; Tadese et al., 2025). Third, studies of geography in Nigeria have worked with states and regions rather than survey clusters, so the specific communities carrying the burden remain unseen (Oladimeji et al., 2026). There is therefore no current, cluster-level account of where unmet need concentrates among AGYW in North-Central Nigeria, and it is not known whether geographic information improves prediction beyond individual and household characteristics.

II. RELATED WORKS

Regional evidence shows that unmet need among young African women is high and socially patterned. Ahinkorah (2020) pooled Demographic and Health Survey data on 24,898 AGYW from ten high-fertility countries, Nigeria among them, and found age, education, occupation, media exposure, parity, wealth, sex of household head, residence and decision-making power to be significant predictors. Teshale (2022) analyzed nineteen surveys and showed that community-level variation in unmet need was substantial and could not be reduced to individual characteristics. Country studies point the same way: in Tanzania, unmet need among sexually active young women was still 27.5% in 2015 to 2016 (Kagoye et al., 2024). In Nigeria, Oyinlola et al. (2024) pooled three survey rounds and found that unmet need among married women fell from 20.2% in 2008 to 16.1% in 2013 before rising to 18.9% in 2018. Odjesa and Okonofua (2024) reported a median odds ratio of 2.245 for demand satisfied by modern methods, larger than most individual-level effects, which indicates that the community a woman lives in matters a great deal. Muhammad et al. (2026) found modern contraceptive use among AGYW to be lower among the less educated, Muslim, rural and northern respondents. Awareness is not the main barrier. A systematic review concluded that uptake has stayed low despite broad knowledge of methods (Ayo-Bello & Quinn-

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

Walker, 2025), and in a facility-based study in southwest Nigeria fear of side effects was the commonest reason for non-use, cited by 66.3% of women with unmet need (Uthman et al., 2022). Psychosocial factors and social support also shape whether sexually active adolescent girls use a method (Musau et al., 2025).

Spatial epidemiology provides formal tools for testing whether an outcome clusters and where Global Moran's I summarizes clustering across a study area, local indicators of spatial association locate the clusters, and the Getis-Ord Gi* statistic identifies statistically significant hotspots. These methods are now applied to Demographic and Health Survey data linked to the displaced cluster coordinates released by the DHS Program (Abate et al., 2025; Burgert et al., 2013). In machine learning, ensemble tree methods gave the best performance for informed contraceptive choice across six high-fertility countries (Melaku et al., 2025), a random forest was the strongest predictor of pregnancy loss (Yehuala et al., 2025), and a stacked ensemble outperformed every individual learner in predicting vaccination dropout in East Africa (Alemayehu et al., 2025). In Nigeria, machine learning has been used to predict fertility preferences (Tadese et al., 2025). Two points emerge. Ensemble methods are the strongest performers on tabular survey data, but none of these applications has studied unmet need among AGYW, and none has added cluster coordinates and spatial dependence to the predictive model. Validation also needs care: observations within a survey cluster are correlated, and cross-validation that ignores such structure gives optimistic estimates of performance (Roberts et al., 2017). The present study addresses both points for North-Central Nigeria using the most recent survey.

III. RESEARCH METHODOLOGY

The study used a cross-sectional design based on the 2023 to 2024 NDHS, a nationally representative survey with a two-stage stratified cluster design (NPC et al., 2025). The women's individual recode was linked to the survey's geographic dataset, which gives the coordinates of each cluster after random displacement of up to 2 km in urban areas and 5 km in rural areas (Burgert et al., 2013); state boundaries for the maps were obtained by dissolving the GRID3 ward boundaries to state level (GRID3, 2020). Of the 2,932 women aged 15 to 24 years in the North-Central zone, those outside the denominator of the revised unmet need indicator (variable v626a) were excluded: women who had never had sex (1,306), unmarried women with no intercourse in the 30 days before interview (247) and infecund women (89) (Bradley et al., 2012). The analytic sample of 1,290 women in 245 clusters comprised 1,003 married or cohabiting, 278 never-married and 9 formerly married women; unmarried women are included only when sexually active, using contraception, pregnant or postpartum amenorrhoeic. The outcome was coded 1 for unmet need for spacing or limiting (codes 1 and 2 of v626a, 296 women) and 0 for current users (codes 3 and 4) and women with no unmet need (code 7), making the task supervised binary classification. Predictors formed three blocks. Individual and household characteristics covered demographic and socioeconomic status, reproductive history (children ever born, age at first sex) and access to information and services. Community context comprised the share with secondary education, the share using contraception, mean wealth and mean exposure to family planning messages among women aged 25 to 49 in the same cluster, so that these measures share no records with the sample. Spatial features were latitude, longitude and a spatial lag, the mean proportion with unmet need in the eight nearest training clusters, excluding the woman's own cluster; it was rebuilt inside every training fold, so held-out outcomes never entered a predictor.

Weighted prevalence was estimated with intervals at the 0.95 level by Taylor linearization over the 14 strata and 262 primary sampling units of the zone (Croft et al., 2018). For spatial analysis,

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

unmet need was aggregated to the 189 clusters with at least three respondents (56 clusters had one or two), with a row-standardized eight-nearest-neighbour weights matrix. Global clustering was tested with Moran's I (999 permutations):

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \times \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \quad (1)$$

where n is the number of clusters, y_i the proportion with unmet need in cluster i, $\bar{y}$ the mean and w_{ij} the spatial weight linking clusters i and j. Hotspots were located with the Getis-Ord Gi* statistic, which includes the focal cluster in its own neighbourhood; S is the standard deviation of y and large positive values mark hotspots:

$$G_i^* = \frac{\sum_j w_{ij} y_j - \bar{y} \sum_j w_{ij}}{S \sqrt{\frac{n \sum_j w_{ij}^2 - \left(\sum_j w_{ij}\right)^2}{n-1}}} \quad (2)$$

Hotspot reliability was checked by false discovery rate control, by varying the number of neighbours (6, 8, 10) and the minimum cluster size (1, 3, 5), and by displacing every cluster a second time 200 times under the DHS rules and recomputing Gi*. Associations with unmet need were screened with the Pearson chi-square test of independence on the two-way table of counts:

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (3)$$

$$E_{ij} = \frac{R_i \times C_j}{N} \quad (4)$$

where O_{ij} is the observed count in row i and column j, R_i and C_j are the row and column totals and N is the sample size. Because this test ignores the survey design, each association was also tested by survey-weighted logistic regression with linearized variance and an adjusted Wald F test on 248 design degrees of freedom.

Eight classifiers were trained: binary logistic regression (equation 5) as the conventional baseline, random forest, extra trees, histogram-based gradient boosting, XGBoost, LightGBM, a soft-voting ensemble of the random forest, XGBoost and LightGBM, and a stacking ensemble in which a logistic meta-learner g combines out-of-fold probabilities $p_1(x)$, $p_2(x)$ and $p_3(x)$ from the same three learners, generated with cluster-grouped folds (equation 3.6). Hyperparameters were fixed at conservative values, and no class weighting was used, so that probabilities stay on the observed scale.

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{j=1}^{k} \beta_j x_j \quad (5)$$

$$P(x) = g(p_1(x), p_2(x), p_3(x)) \quad (6)$$

Models were compared by five repeats of stratified cluster-grouped five-fold cross-validation on the full sample, which was used to select the final model (Roberts et al., 2017). A held-out test set of 49 clusters (298 women, 61 with unmet need) gave an independent check, with intervals from 2,000 cluster bootstrap resamples. Each model's decision threshold was the value maximizing F1 on out-of-fold training predictions, fixed before the test set was used. Discrimination was measured by the area under the receiver operating characteristic curve (ROC-AUC) and the precision-recall curve (PR-AUC); calibration by the Brier score and the calibration slope and intercept; and classification by precision, recall and F1, where TP, FP and FN denote true positives, false positives and false negatives:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

A sensitivity analysis refitted every model with the normalized survey weights. Predictors were interpreted by permutation importance on the final model and by Shapley values computed exactly for LightGBM (Lundberg et al., 2020); both describe predictive associations, not causal effects. The contribution of geography was measured by an ablation in which five models were trained on individual and household features, then with community context, then with spatial features, within each cross-validation repeat.

IV. RESULTS AND DISCUSSION

The results obtained in this study is discussed in this section:

A. Data Presentation

Most of the 1,290 young women were aged 20 to 24 years (73.3%) and married or cohabiting (80.0%); 44.2% had no formal education and 74.3% lived in rural areas. The weighted prevalence of unmet need was 22.1% (interval 19.3% to 25.0%). Figure 1 shows the state values as a map and as a bar chart. Prevalence ranged from 35.8% in Kwara and 30.7% in Kogi, two neighbouring states in the south-west of the zone, to 17.8% in Plateau and 17.4% in Niger.

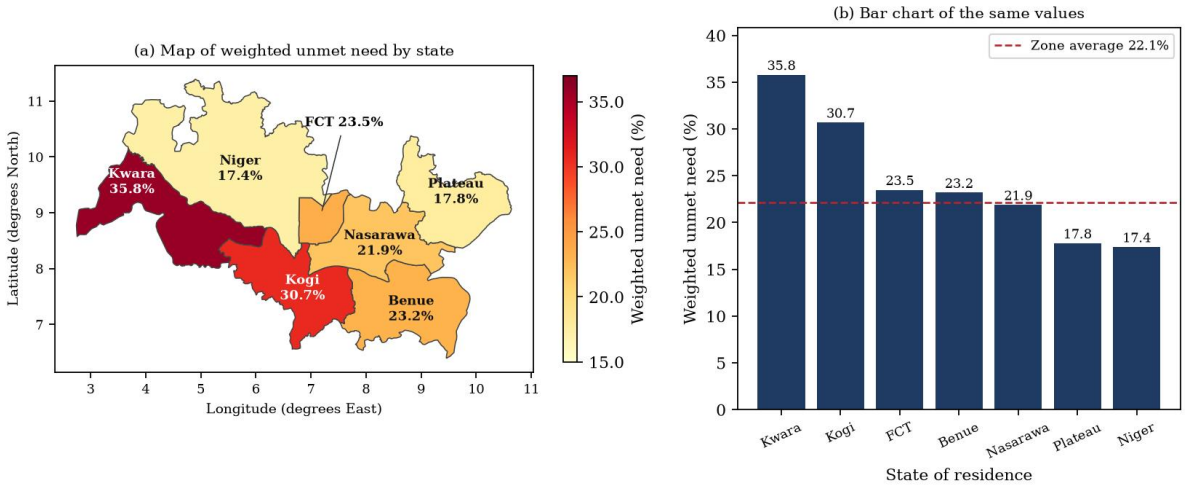


Figure 1: Weighted Prevalence of Unmet need by State

Table 1 shows why the two extremes differ. In Kwara, 34.4% of the at-risk young women had never married and 70.5% of these had an unmet need; they made up 35 of the 52 Kwara women with unmet need. Among married young women, unmet need in Kwara (17.5%) was close to that in Niger (15.9%), so the Kwara figure is driven by sexually active young women outside marriage who use no method. In Niger, 97.7% of the at-risk sample were married, 85.7% had no education, only 5.3% were using any method and none reported a visit from a family planning worker; 77.2% wanted a child soon and so, under the survey definition, had no unmet need. Low unmet need in Niger therefore reflects early marriage and early childbearing, not better services. Allowing for the survey design, the intervals for the two states, 25.8% to 45.8% and 12.5% to 22.3%, do not overlap.

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

Table 1: Unmet Need and Selected Characteristics of the At-risk Sample by State (weighted %)

State	n	Unmet need	Never married	Unmet need, never married	Unmet need, married	Using a method	No education
Kwara	150	35.8	34.4	70.5	17.5	12.1	41.4
Kogi	173	30.7	23.5	53.9	24.3	17.5	27.9
FCT	65	23.5	34.8	37.3	16.2	31.7	13.4
Benue	192	23.2	30.9	22.9	21.5	36.6	6.7
Nasarawa	187	21.9	19.9	25.8	20.9	24.7	43.4
Plateau	269	17.8	25.3	30.0	13.1	32.9	20.4
Niger	254	17.4	2.3	86.3[a]	15.9	5.3	85.7
North-Central	**1,290**	**22.1**	**19.3**	**39.7**	**17.6**	**20.3**	**44.2**

[a] Based on 7 never-married women; too few for a reliable estimate. Married excludes the 9 formerly married women.

Unmet need was positively spatially autocorrelated (Moran's I = 0.0797, expected −0.0053, z = 2.461, p = 0.013). Local analysis classified 12 clusters as high-high, all in Kwara, and 9 as low-low. Gi* identified 15 hotspots and 16 coldspots (Figure 2a): 13 hotspots lie in Kwara, with 1 in FCT and 1 in Kogi, and the coldspots fall mainly in Nasarawa (8) and Niger (5). The concentration in Kwara proved robust. After false discovery rate control 8 hotspots remained, all in Kwara; with at least three women per cluster the count ranged from 14 to 16 across neighbourhood sizes, with Kwara holding the largest share in every setting; and across 200 re-displacements the median count was 13, with 12 of the 15 original hotspots significant in at least 80.0% of simulations (Figure 2b). The intra-class correlation coefficient was 10.0% and the median odds ratio 1.78.

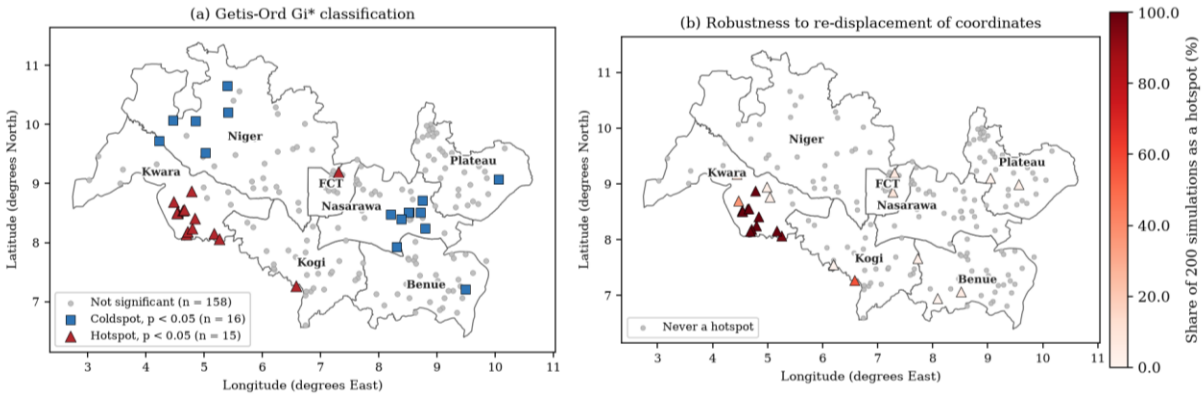


Figure 2: Getis-Ord Gi* Hotspots (a) and their Robustness to Repeated Displacement of Coordinates (b)

Table 2 reports the bivariate tests. Four characteristics were significant by the Pearson test: educational level, marital or union status, sex of household head and state of residence. For state, Kwara alone contributes 11.66 of the χ^2 of 22.01, with 52 women with unmet need against 34.42 expected under independence. After allowing for the survey design, educational level, marital or union status and state of residence remained significant, whereas sex of household head did not.

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

Table 2: Pearson and Design-adjusted Tests of Association with Unmet Need

Characteristic	χ^2	df	p	Design F	Design p	Design remark
Age group	2.63	1	0.105	2.09	0.150	Not significant
Educational level	**20.04**	**3**	**<0.001**	**4.69**	**0.003**	**Significant**
Religion	5.04	3	0.169	1.39	0.245	Not significant
Marital or union status	**62.48**	**2**	**<0.001**	**18.49**	**<0.001**	**Significant**
Place of residence	3.48	1	0.062	0.79	0.376	Not significant
Wealth quintile	6.56	4	0.161	1.55	0.189	Not significant
Working status	3.51	1	0.061	1.48	0.225	Not significant
Sex of household head	7.02	1	0.008	3.36	0.068	Not significant
State of residence	**22.01**	**6**	**0.001**	**2.76**	**0.013**	**Significant**
Knows a modern method	0.13	1	0.716	0.06	0.803	Not significant
Internet use	2.01	1	0.157	1.41	0.236	Not significant
Health insurance cover	0.30	1	0.585	1.45	0.230	Not significant
Visited by family planning worker	3.62	1	0.057	3.65	0.057	Not significant
Visited a health facility	0.03	1	0.866	0.00	0.959	Not significant
Distance to facility a big problem	0.55	1	0.457	0.07	0.792	Not significant
Awareness of STI	3.61	1	0.057	2.00	0.158	Not significant

Table 3 summarizes model performance. Under repeated cross-validation the machine learning models reached a mean ROC-AUC of 0.684 to 0.713, against 0.637 for logistic regression; the voting ensemble was highest (0.713, SD 0.008), but the ensembles differed by about 0.01, less than the variation between repeats. On the test set the voting ensemble reached 0.698 (0.610 to 0.780); at its training-derived threshold of 0.22 it identified 63.9% of young women with unmet need, with precision 0.342. The stacking-minus-baseline difference in test ROC-AUC was 0.135 (0.039 to 0.225), so the null hypotheses of no difference among models and of no gain over the baseline are rejected. Training ROC-AUC far exceeded the cross-validated values, as expected of flexible learners, but the honest estimates agreed closely. Refitting with survey weights changed test ROC-AUC by at most 0.023. A stacking ensemble built with the default scikit-learn routine, whose internal folds follow the file order and ignore clusters, scored only 0.288.

Table 3: Training, Repeated Cross-validation and Test Performance of the Models

Model	Train ROC-AUC	CV ROC-AUC (SD)	Test ROC-AUC (interval)	Test F1	Brier	Calib. slope
Logistic regression	0.739	0.637 (0.004)	0.565 (0.473 to 0.666)	0.339	0.168	0.33
Random Forest	0.961	0.707 (0.008)	0.698 (0.614 to 0.781)	0.435	0.150	1.39
Extra Trees	0.963	0.684 (0.010)	0.670 (0.591 to 0.750)	0.361	0.150	1.16
Gradient Boosting	0.951	0.704 (0.010)	0.678 (0.583 to 0.763)	0.417	0.160	0.56
XGBoost	0.968	0.703 (0.006)	0.701 (0.617 to 0.778)	0.361	0.158	0.60
LightGBM	0.962	0.703 (0.011)	0.670 (0.582 to 0.748)	0.415	0.165	0.49
Voting Ensemble	0.969	0.713 (0.008)	0.698 (0.610 to 0.780)	0.446	0.153	0.81
Stacking Ensemble	0.969	0.710 (0.008)	0.700 (0.612 to 0.784)	0.425	0.153	1.03

CV: five repeats of cluster-grouped five-fold cross-validation on the full sample. Test intervals are 2.5th to 97.5th percentiles of 2,000 cluster bootstrap resamples; F1 uses each model's training-derived threshold. A calibration slope of 1 indicates good calibration.

Figure 3 compares observed and predicted risk on the test set. The stacking ensemble, whose logistic meta-learner recalibrates the base predictions, was closest to the diagonal (slope 1.03, Brier 0.153). The voting ensemble was slightly over-confident (slope 0.81), the random forest too

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

cautious (slope 1.39), the boosting learners over-confident, and logistic regression badly so (slope 0.33).

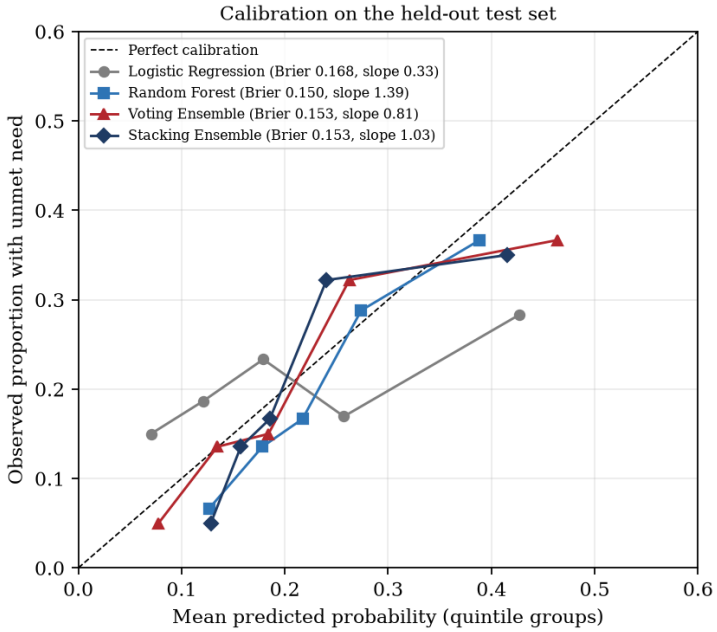


Figure 3: Calibration of Selected Models on the Held-out Test Set

Permutation importance on the voting ensemble ranked marital or union status (0.127) and children ever born (0.079) far ahead of latitude, household wealth index and longitude. Shapley values for LightGBM placed the same two predictors first: being never married sharply raised predicted risk and having no children lowered it. The coordinates rank highly because the fitted model uses them in place of information that state and community composition also carry. Table 4 gives the ablation. Community context added +0.006 to +0.019 to the mean ROC-AUC and the spatial variables a further 0.000 to +0.008, small and inconsistent increments, so location adds little that the other predictors do not already supply. When the spatial lag was instead built from all clusters before splitting, it appeared to add 0.045 for a random forest, because held-out outcomes had leaked into the predictor.

Table 4: Mean ROC-AUC by Feature Set under Repeated Cross-Validation

Model	Individual + household	+ Community	+ Spatial	Spatial gain	Range over repeats
Logistic regression	0.628	0.633	0.637	+0.004	0.000 to +0.010
Random Forest	0.686	0.699	0.707	+0.008	−0.007 to +0.013
XGBoost	0.687	0.703	0.703	0.000	−0.006 to +0.006
LightGBM	0.687	0.699	0.703	+0.005	−0.013 to +0.013
Stacking Ensemble	0.687	0.706	0.710	+0.004	0.000 to +0.008

B. Discussion of Findings

The prevalence of 22.1% is higher than the 18.9% reported for married Nigerian women in 2018 (Oyinlola et al., 2024), largely because the denominator here is restricted to young women exposed to pregnancy, and it is close to the 27.5% reported for young Tanzanian women (Kagoye et al., 2024). Unmet need is concentrated at the scale of groups of neighbouring clusters within a single zone, which extends the administrative-level evidence of Oladimeji et al. (2026) and Sanni et al.

https://doi.org/10.5281/zenodo.24199907

(2025). Because the coordinates are displaced and some hotspot clusters at the edge of the concentration vary with the analytic choices, the result identifies a high-need area in southern Kwara rather than exact settlements.

The state contrast shows why unmet need must be read with care. Because the indicator is the gap between the wish to delay a birth and contraceptive use (Bradley et al., 2012), it can be low either because services meet demand or because the demand to delay is itself low. Niger is the second case, consistent with the low uptake reported for young women in northern Nigeria (Muhammad et al., 2026; Sanni et al., 2025). Kwara reflects a different gap, among unmarried young women who want to avoid pregnancy but use no method, a group whose contraceptive behaviour depends on psychosocial factors and social support as well as knowledge (Musau et al., 2025).

The leading predictors agree with Ahinkorah (2020) and Muhammad et al. (2026). The ensembles outperformed regression, as in regional work (Melaku et al., 2025; Yehuala et al., 2025), but the discrimination reached is moderate, and voting and stacking performed alike, so a learned combination added little over a simple average (Alemayehu et al., 2025). Location added little to individual prediction even though unmet need clusters in space: the clustering is largely explained by the characteristics of the women and communities in the high-need area, which the models already contain. Two pitfalls emerged, a stack whose internal folds ignored clusters and a spatial lag built before splitting, and both show that folds and outcome-based features must respect the clustered design (Roberts et al., 2017).

V. CONCLUSION

Unmet need for contraception affects 22.1% of at-risk adolescent girls and young women in North-Central Nigeria and is geographically concentrated, above all in southern Kwara, a pattern that survived correction for multiple testing and repeated displacement of the coordinates. All four null hypotheses were rejected. Ensemble learners predicted unmet need better than logistic regression (mean ROC-AUC about 0.71 against 0.64), and stacking gave well-calibrated probabilities, but location added little once individual, household and community characteristics were known. The predictors identify who is at higher risk; they are associations, not causes. The study is limited by its cross-sectional design, by coordinate displacement, which restricts the spatial results to the scale of groups of communities, and by modest state samples. This study recommendations the following:

(a). Family planning programmes in the zone should prioritize the high-need area of southern Kwara, treating it as a priority zone at the scale of neighbouring communities or local government areas and confirming the specific communities with facility and community data before allocating resources.

(b). Services and communication in Kwara and Kogi should be designed for sexually active young women who have never married, with confidential and youth-friendly provision.

(c). In Niger, low measured unmet need should not be taken as a sign of adequate services; outreach and birth-spacing counselling for young married women and their partners are needed.

(d). Risk scores of the kind developed here can help direct further assessment, but their moderate discrimination means they should not decide individual eligibility, and interventions suggested by the predictors should be evaluated in their own right.

(e). Researchers applying machine learning to survey data should use cluster-grouped validation and build outcome-based spatial features inside the validation folds; future work should extend the analysis to all six zones and to successive survey rounds.

Volume 2, Issue 4, 2026
https://doi.org/10.5281/zenodo.24199907

REFERENCES

Abate, B. J., Melesse, A. W., Brhan, H., & Agimas, M. C. (2025). Spatial variation, pooled prevalence, and factors associated with perinatal mortality in sub-Saharan Africa, evidence from demographic and health surveys 2015–2023: A geospatial regression approach. eClinicalMedicine, 81, 103137. https://doi.org/10.1016/j.eclinm.2025.103137

Ahinkorah, B. O. (2020). Predictors of unmet need for contraception among adolescent girls and young women in selected high fertility countries in sub-Saharan Africa: A multilevel mixed effects analysis. PLOS ONE, 15(8), e0236352. https://doi.org/10.1371/journal.pone.0236352

Alemayehu, M. A., Kebede, S. D., Walle, A. D., Mamo, D. N., Enyew, E. B., & Adem, J. B. (2025). A stacked ensemble machine learning model for the prediction of pentavalent 3 vaccination dropout in East Africa. Frontiers in Big Data, 8, 1522578. https://doi.org/10.3389/fdata.2025.1522578

Ayo-Bello, O., & Quinn-Walker, N. (2025). Factors associated with modern contraceptives use in Nigerian women: A systematic review of quantitative studies. Discover Public Health, 22, 300. https://doi.org/10.1186/s12982-025-00674-8

Bradley, S. E. K., Croft, T. N., Fishel, J. D., & Westoff, C. F. (2012). Revising unmet need for family planning (DHS Analytical Studies No. 25). ICF International.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys (DHS Spatial Analysis Reports No. 7). ICF International.

GRID3. (2020). GRID3 Nigeria operational ward boundaries [Data set]. Geo-Referenced Infrastructure and Demographic Data for Development. https://data.grid3.org

Kagoye, S. A., Jahanpour, O., Ngoda, O. A., Obure, J., Mahande, M. J., & Renju, J. (2024). Trends and determinants of unmet need for modern contraception among adolescent girls and young women in Tanzania, 2004–2016. PLOS Global Public Health, 4(1), e0000695. https://doi.org/10.1371/journal.pgph.0000695

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Melaku, M. S., Yohannes, L., Sharew, B., Derseh, M. H., & Taye, E. A. (2025). Application of machine learning algorithms to model predictors of informed contraceptive choice among reproductive age women in six high fertility rate sub-Saharan Africa countries. BMC Public Health, 25, 1986. https://doi.org/10.1186/s12889-025-23242-w

Muhammad, A., Okunlola, M., & Morhason-Bello, I. (2026). Determinants of modern contraceptive use among adolescent girls and young women in Nigeria. International Journal of Community Medicine and Public Health, 13(7), 3359–3370. https://doi.org/10.18203/2394-6040.ijcmph20262225

Musau, A., Mtei, E., Ogbondeminu, F., & Lundgren, R. (2025). Psychosocial determinants of contraceptive desire and use among sexually active adolescent girls in Kenya and Nigeria: Implications for girl-centered contraceptive programs. Contraception and Reproductive Medicine, 10, 118. https://doi.org/10.1186/s40834-025-00389-w

National Population Commission (NPC) [Nigeria], Federal Ministry of Health and Social Welfare, & ICF. (2025). Nigeria Demographic and Health Survey 2023–24: Final report. NPC and ICF. https://www.dhsprogram.com/pubs/pdf/FR395/FR395.pdf

Odjesa, E., & Okonofua, F. E. (2024). An empirical analysis of the demand for family planning satisfied by modern methods among married or in-union women in Nigeria: Application of multilevel binomial logistic modelling technique. PLOS ONE, 19(3), e0300744. https://doi.org/10.1371/journal.pone.0300744

Oladimeji, D., Zaman, J. N., & Kwaghe, V. (2026). Persistent spatial inequalities in contraceptive use in Nigeria: A Bayesian spatio-temporal model, evidence from Demographic and Health Survey data from 2003 to 2024 [Preprint]. SSRN. https://doi.org/10.2139/ssrn.6904561

Oyinlola, F. F., Kupoluyi, J. A., & Adetutu, O. M. (2024). Changes in unmet need for family planning among married women of reproductive age in Nigeria: A multilevel analysis of a ten-year DHS wave. PLOS ONE, 19(8), e0306768. https://doi.org/10.1371/journal.pone.0306768

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929. https://doi.org/10.1111/ecog.02881

Sanni, O. F., Sanni, A. E., Oyedeji, A. O., Sanni, R. O., & Abiodun, O. (2025). A comparative analysis of factors associated with modern contraceptive use among youth in Northern and Southern Nigeria: A cross-sectional population-based survey (2011–2021). PLOS Global Public Health, 5(6), e0004685. https://doi.org/10.1371/journal.pgph.0004685

Solanke, B. L., Adetutu, O. M., Rahman, S. A., Soladoye, D. A., & Owoeye, M. O. (2022). Prevalence and determinants of unmet need for contraception among women in low and high-priority segments for family planning demand generation in Nigeria. Archives of Public Health, 80, 245. https://doi.org/10.1186/s13690-022-00997-x

Tadese, Z. B., Nimani, T. D., Mare, K. U., Gubena, F., Wali, I. G., & Sani, J. (2025). Exploring machine learning algorithms for predicting fertility preferences among reproductive age women in Nigeria. Frontiers in Digital Health, 6, 1495382. https://doi.org/10.3389/fdgth.2024.1495382

Teshale, A. B. (2022). Factors associated with unmet need for family planning in sub-Saharan Africa: A multilevel multinomial logistic regression analysis. PLOS ONE, 17(2), e0263885. https://doi.org/10.1371/journal.pone.0263885

Uthman, M. K., Bello, I. S., Fadugbagbe, A. O., Olajubu, T. O., Ismail, W. O., & Ibrahim, A. O. (2022). Unmet needs for family planning and its determinants among women of reproductive age in Ilesha, Southwest Nigeria: A cross-sectional study. Journal of Medicine Access, 6, 1–10. https://doi.org/10.1177/27550834221115979

Yehuala, T. Z., Mengesha, S. B., & Baykemagn, N. D. (2025). Predicting pregnancy loss and its determinants among reproductive-aged women using supervised machine learning algorithms in sub-Saharan Africa. Frontiers in Global Women's Health, 6, 1456238. https://doi.org/10.3389/fgwh.2025.1456238