

Predictors of fertility-limiting preferences among currently married women in Bangladesh: an interpretable machine learning analysis of the 2022 Demographic and Health Survey

Md. Mahbub-Ur Rashid

Mawlana Bhashani Science and Technology University

Prayas Sur Antu

prayassurant@gmail.com

Mawlana Bhashani Science and Technology University

Ujjoyene Talapatra Athai

Mawlana Bhashani Science and Technology University

Md. Hazrat Ali

Mawlana Bhashani Science and Technology University

Research Article

Keywords: Fertility preference, Family planning, Machine learning, SHAP (Shapley Additive exPlanations), Internal migration, Demographic and Health Survey, Bangladesh

Posted Date: October 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10977003/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Bangladesh's total fertility rate has stalled at 2.3 births per woman since 2011. Knowing which women want to stop childbearing helps family planning programmes direct limiting and spacing services. To support these programmes, this study aimed to predict fertility-limiting preferences among currently married women in Bangladesh and to identify the most influential predictors, including internal migration, digital inclusion and mental health.

Methods

Data on 17,245 currently married women aged 15–49 years from the 2022 Bangladesh Demographic and Health Survey were analyzed. Twenty-five candidate predictors, represented by 38 encoded features, were screened with Boruta on an 80% training set. Five classifiers were tuned by 10-fold stratified cross-validation and evaluated on a 20% held-out test set, and the best-performing model was interpreted using Shapley additive explanations (SHAP).

Results

Overall, 60.9% of the women wanted no more children. XGBoost performed best, with a test area under the curve (AUC) of 0.926 (95% CI 0.917–0.934), sensitivity of 84.1% and specificity of 84.3%. Age group, number of living children, contraceptive intention, long-term residence, ideal family size, age at marriage and household size were the most influential predictors. Recent migration, higher education, a larger ideal family size and an intention to begin contraception shifted predictions toward a non-limiting preference, whereas digital inclusion and anxiety or depression contributed relatively little to model predictions.

Conclusions

Fertility-limiting preferences can be accurately predicted from a few routinely collected characteristics. The duration of residence and contraceptive intention could help programmes identify women who want to stop childbearing for long-acting methods and recent migrants for spacing services.

Background

A woman's stated fertility preference, whether she wants another child or wants no more, is one of the strongest predictors of her subsequent contraceptive behavior and childbearing [1], and its socioeconomic and demographic correlates have been widely documented [2, 3]. Preferences indicate programmes that women need contraception to stop childbearing and that need it to delay the next birth; when preferences are not matched by contraceptive practice, unmet needs follow [4, 5], and unintended pregnancy remains common in low- and middle-income settings [6]. Matching services to what women want is therefore central

to rights-based family planning and to the Sustainable Development Goal target of universal access to sexual and reproductive health services [7, 8].

Bangladesh, one of the world's most populous countries, illustrates why this matters. Its total fertility rate fell from more than six births per woman in the 1970s to 2.3 by 2011, largely through a sustained national family planning programme, but it has not moved since, remaining at 2.3 in 2022 and above the replacement level [9, 10]. Approximately three in five currently married women say that they want no more children, yet an unmet need for family planning persists, child marriage remains common, and son preference and regional differentials, which have long been documented in Bangladesh [11] and elsewhere in South Asia [12], continue to shape reproductive behavior [9]. In such a setting, the pressing question is no longer why fertility has fallen but which women hold a limiting preference so that those who want to stop can be reached with long-acting methods and those who want to space can be reached with appropriate services.

Most studies of fertility preference in sub-Saharan Africa have used logistic regression [2, 3, 5]. These studies consistently identify age, parity, education, wealth, religion and contraceptive use as correlates, but regression assumes that each factor acts additively on the log-odds, whereas the meaning of a given parity depends on a woman's age, and education modifies both. Machine learning (ML) algorithms relax these assumptions and match, and in some tasks, they outperform, conventional models in terms of epidemiological and clinical prediction [13, 14, 15]. Applications to fertility outcomes in low- and middle-income countries are recent. Tadese et al. compared six classifiers among 37,581 women in the 2018 Nigeria Demographic and Health Survey (DHS) and reported that tree ensembles discriminated best, with parity, age and ideal family size among the leading predictors [16]. Sani et al. combined random forests with Shapley additive explanations (SHAPs) among 8,951 Somali women, reaching an accuracy of 81%, with age, region and recent births dominating the explanations [17]. Related work has classified fertility levels in Ethiopia [18], predicted childbearing intentions among Iranian women [19] and, in Bangladesh, predicted short birth intervals from the 2022 DHS [20]. These studies share three limitations. All nationally representative, DHS-based fertility-preference applications were conducted in high-fertility African settings, where pronatalist norms keep preferences high across the life course, whereas one late-transition application [19] used a small clinic-recruited sample of women approaching marriage and applied no interpretability framework; they relied on the standard sociodemographic predictors available in every DHS; and class imbalance was handled with synthetic oversampling [17], which can distort the joint distribution of categorical survey data and inflate reported performance.

The gap is also substantive because several influences on fertility preference that are prominent in Bangladesh have rarely been measured together. Internal migration influences childbearing through selection, adaptation and disruption [21], and rural-to-urban movement is reshaping reproductive norms across the country; however, migration is usually reduced to an urban–rural indicator. Mobile phone diffusion has been linked to greater contraceptive use and better maternal and child health [22]. Child loss can sustain replacement motives for continued childbearing [23], and longitudinal evidence from Norway and Sweden indicates that depression is negatively associated with fertility intentions and their realization among women [24]. The 2022 Bangladesh DHS (BDHS) allows these factors to be examined together for the first

time: it is the first round to include anxiety [25] and depression [26] screening modules [9], and its residence history and phone-and-internet items permit a graded migration classification and a digital inclusion index.

This study aimed to predict fertility-limiting preferences among currently married women aged 15–49 years in the 2022 BDHS and to explain which characteristics drive the prediction. Three questions were addressed: how accurately can fertility-limiting preferences be predicted from survey-measured characteristics; which predictors, among sociodemographic, reproductive, migration, digital inclusion and mental health domains, carry contribute most strongly to the model's predictions; and in what direction does each predictor act? Boruta all-relevant feature selection [27] was combined with five classifiers spanning linear, kernel, bagging and gradient-boosting families, and the best model was interpreted with SHAP [28]. To our knowledge, no previous study has applied an explainable ML framework to fertility preferences in Bangladesh or in another late-transition South Asian population, where the question is who wants to stop childbearing rather than why fertility remains high; extends the predictor set to migration duration, digital inclusion, autonomy, attitudes toward intimate partner violence, child loss and screened anxiety and depression; and identifies routinely collected markers that programmes could use to distinguish women who want to stop childbearing from those who want to space.

Methods

Study design and data sources

This cross-sectional study used secondary data from the 2022 BDHS, a nationally representative household survey implemented by the National Institute of Population Research and Training (NIPORT) with technical assistance from the ICF through the DHS Program [7]. The survey used a two-stage stratified cluster design: 675 enumeration areas (237 urban areas and 438 rural areas) were selected with a probability proportional to size, and 45 households were systematically selected within each. Field work took place from 27 June to 12 December 2022, and 30,078 ever-married women aged 15–49 years were interviewed (response rate 99.1%). In two-thirds of households, women received a long questionnaire, which included modules on husbands' characteristics and women's work and mental health; women in the remaining third received a short questionnaire covering background characteristics and reproductive history only [7]. The individual record file (BDIR81SV) was analyzed.

Study population and sample selection

The analytical sample was restricted to currently married women who completed the long questionnaire and had a classifiable fertility preference. Among the 30,078 ever-married women, 1,541 who were widowed, divorced, separated or living with a partner were excluded, as were the 9,550 women interviewed with the short questionnaire, which did not include the mental health module. A further 686 women who declared themselves infecund were excluded because the fertility-preference question does not apply to them, and 933 visitors to the household were excluded because the duration of residence, and hence migration status, could not be determined. After listwise deletion of 123 women with incomplete covariate data (74 with a

nonnumeric ideal family size, 15 with incomplete anxiety items and 38 with incomplete depression items, some overlapping), the final sample comprised 17,245 women (Additional file: Fig. S1).

Outcome variables

Outcome variables

The outcome was derived from the DHS question on the desire for more children. Women who wanted no more children or who were sterilized or whose husband was sterilized were classified as having a limiting preference (coded 1). Women who wanted another child within two years, after two or more years or at an undecided time, and women who were undecided about wanting another child were classified as nonlimiting (coded 0). Sterilization was treated as a revealed limiting preference; declared infecundity was treated as a constraint rather than a preference. The limiting category was the positive class in all model training, evaluation and interpretation.

Explanatory variables

Twenty-five predictors were selected on the basis of previous research and availability in the BDHS [2, 16, 17]. Sociodemographic characteristics included age group (15–24, 25–34, 35–49 years), division, place of residence, religion, household size (1–3, 4–6, ≥ 7 members), sex of the household head and wealth index (poor, middle, rich). The human capital and economic variables were the respondent's and husband's education (none, primary, secondary, higher), the husband's occupation and the respondent's working status. The reproductive and marital variables used were the number of living children (0, 1–3, ≥ 4), current pregnancy, experience of child death, age at first marriage (< 15 , 15–17, ≥ 18 years), ideal family size (0–2 vs ≥ 3), and the husband's coresidence and contraceptive behavior (current user; nonuser intending to use; nonuser with no intention; nonuser unsure or other). Migration status was derived from years lived in the current locality, following the duration-based approach used in the migration–fertility literature [21]: always resident, resident for 15 or more years, established migrant (5–14 years) and recent migrant (0–4 years); the cutoff points were defined by the authors to separate recent arrivals from settled residents. A digital-inclusion index summed internet use in the previous 12 months, smartphone ownership, holding a bank or mobile-money account and having used it in the previous 12 months (low 0, medium 1–2, high 3–4); items skipped because a woman owned no phone or no account were scored as absent rather than missing. Media exposure was any reported use of newspaper, radio or television. The psychosocial variables were autonomy in household decision-making (the number of three decisions made alone or jointly: low 0, medium 1–2, high 3), whether any of the five reasons for wife beating was considered justified, anxiety severity from the GAD-7 (minimal 0–4, mild 5–9, moderate/severe 10–21) [25] and depression severity from the PHQ-9 (minimal 0–4, mild 5–9, moderate/severe 10–27) [26]. "Do not know" and "refused" responses to the GAD-7 and PHQ-9 items were set to missing, and a scale was scored only when all items were answered.

Statistical analysis

Following DHS conventions, counts are unweighted, and percentages are weighted with individual sampling weights to account for the complex survey design [9]. Associations between each predictor and fertility

preference were assessed with Pearson chi-square tests on weighted contingency tables in which weights were normalized to sum to the sample size so that the effective sample size was not inflated; no continuity correction was applied. Because the large sample makes most associations statistically significant, Cramér's V was reported as an effect size and interpreted using conventional thresholds as negligible (< 0.10), weak ($0.10-0.29$), moderate ($0.30-0.49$) or strong (≥ 0.50). Statistical significance was set at two-sided $p < 0.05$.

Machine learning pipeline

The prediction model was reported in accordance with the TRIPOD + AI statement [29]. To prevent information leakage, the sample was divided into a training set (80%; $n = 13,796$) and a held-out test set (20%; $n = 3,449$), stratified by the outcome, before encoding, feature selection or model fitting [30]. The sampling weight was not used as a predictor. Ordered variables were represented using order-preserving integer codes, whereas nominal variables were one-hot encoded with the first category as the reference; category definitions were determined from the training set and then applied unchanged to the test set, yielding 38 encoded features. All-relevant feature selection was performed with Boruta [27] on the training set only. Boruta compared each encoded feature with randomly permuted shadow features using a random forest base learner with balanced class weights, maximum depth 5 and 100 iterations. The features confirmed by Boruta were subsequently used in model tuning and cross-validation. Because feature selection was performed once before cross-validation rather than independently within each fold, the cross-validated AUCs may be mildly optimistic; however, the held-out test set remained untouched throughout feature selection, model tuning and final evaluation. The multicollinearity among the selected features was checked with variance inflation factors (VIFs).

Five classifiers representing different learning paradigms were trained on the Boruta-confirmed features: L2-regularized logistic regression, support vector machine (SVM) [31] with linear and radial-basis kernels, random forest [32], XGBoost [33] and CatBoost [34]. The hyperparameters were tuned via an exhaustive grid search under 10-fold stratified cross-validation on the training set, with the area under the receiver operating characteristic curve (AUC) as the selection criterion (**Supplementary Table S1**). For logistic regression and SVM, features were standardized within each cross-validation fold. Class imbalance (61.3% limiting in the training set) was handled with balanced class weights (logistic regression, SVM, random forest and CatBoost) or `scale_pos_weight` (XGBoost) rather than synthetic oversampling, which can distort the joint distribution of categorical survey data. Each tuned model was evaluated once on the test set via the AUC with a 1,000-resample bootstrap 95% confidence interval (CI), the area under the precision–recall curve (PR-AUC), accuracy, precision, recall (sensitivity), specificity and F1 score for the limiting class. The final model was selected on cross-validated AUC.

Model interpretation

The final model was interpreted with SHAP [28] via the exact TreeExplainer algorithm [35, 36]. SHAP values, which decompose each prediction into additive feature contributions on the log-odds scale, were computed on the test set for the limiting class. Global importance was summarized as the mean absolute SHAP value,

and the beeswarm plot was restricted to the smallest set of features accounting for at least 90% of the total. Local explanations were produced for the women with the highest and lowest predicted probabilities.

Analyses were performed in Python 3.13 with pandas 3.0, scikit-learn 1.9 [37], Boruta 0.4.3, XGBoost 3.4, CatBoost 1.2 and SHAP 0.52, with a fixed random seed.

Results

Sample characteristics

Among the 17,245 women, 10,566 (60.9%) wanted no more children, and 6,679 (39.1%) wanted another child or were undecided (Table 1). One quarter were aged 15–24 years (25.0%), 36.5% were aged 25–34 years, 38.5% were aged 35–49 years, 71.4% lived in rural areas, and 90.1% were Muslim. Nearly half of the respondents had secondary education (47.8%), and 12.5% had no secondary education. One in six women (16.5%) had lived in their current locality, 34.2% had lived there for 15 years or longer, 29.8% were established migrants, and 19.6% were recent migrants. Digital inclusion was low for 55.8% of women. Most had one to three living children (79.7%), 10.8% had lost a child, and 67.0% were using contraception. Two-thirds had married before age 18 (27.9% before 15 and 39.9% at 15–17 years). Mild or worse anxiety symptoms were reported by 25.5% of the participants, and depressive symptoms were reported by 28.0%.

Bivariate associations

All the predictors except ideal family size were associated with fertility preference at $p < 0.05$, but the effect sizes varied widely (Table 1). The age group ($V = 0.588$) was strongly associated: the percentage of women who wanted no more children rose from 17.5% among women aged 15–24 years to 91.0% among those aged 35–49 years. Migration status ($V = 0.490$), number of living children ($V = 0.454$) and respondent education ($V = 0.312$) showed moderate associations: 23.6% of recent migrants wanted no more children than did 89.5% of residents aged 15 years or longer; 2.0% of childless women compared with 97.9% of those with four or more children; and 39.0% of women with higher education than did those with no children; and 88.4% of those with no children. Contraceptive behavior, husband's education, age at marriage, digital inclusion, household size, husband's occupation, child loss, husband's coresidence, working status, autonomy and current pregnancy were weakly associated ($V 0.10$ – 0.25). Division, residence, religion, household headship, media exposure, wealth, attitudes toward wife beating and anxiety and depression severity were statistically significant but of negligible magnitude ($V < 0.10$). Ideal family size was not associated with preference ($p = 0.164$).

Table 1
Characteristics of currently married women and bivariate associations with fertility-limiting preferences,
Bangladesh DHS 2022 (n = 17,245)

Characteristic	Category	n (%)	Limiting, %	χ^2 (df)	p	Cramér's V
Age group (years)	15–24	4,208 (25.0)	17.5	5953.31 (2)	< 0.001	0.588
	25–34	6,314 (36.5)	59.0			
	35–49	6,723 (38.5)	91.0			
Division	Barishal	1,816 (5.9)	61.5	64.13 (7)	< 0.001	0.061
	Chattogram	2,570 (18.6)	57.6			
	Dhaka	2,662 (25.7)	58.3			
	Khulna	2,237 (11.9)	65.3			
	Mymensingh	1,837 (7.5)	61.5			
	Rajshahi	2,219 (13.2)	65.1			
	Rangpur	2,117 (11.7)	62.5			
	Sylhet	1,787 (5.5)	59.3			
Place of residence	Urban	6,062 (28.6)	59.3	7.33 (1)	0.007	0.021
	Rural	11,183 (71.4)	61.5			
Migration status	Always resident	2,822 (16.5)	60.4	4140.92 (3)	< 0.001	0.490
	Resident 15 + years	5,910 (34.2)	89.5			
	Established migrant (5– 14 years)	5,175 (29.8)	52.8			
	Recent migrant (0–4 years)	3,338 (19.6)	23.6			
Respondent's education	No education	2,136 (12.5)	88.4	1673.68 (3)	< 0.001	0.312

Characteristic	Category	n (%)	Limiting, %	χ^2 (df)	p	Cramér's V
	Primary	4,513 (26.2)	73.5			
	Secondary	8,057 (47.8)	53.0			
	Higher	2,539 (13.5)	39.0			
Religion	Muslim	15,402 (90.1)	60.3	22.10 (1)	< 0.001	0.036
	Non-Muslim	1,843 (9.9)	66.2			
Household size	Small (1–3)	3,810 (23.1)	49.0	521.88 (2)	< 0.001	0.174
	Medium (4–6)	10,237 (59.2)	67.9			
	Large (≥ 7)	3,198 (17.7)	53.1			
Sex of household head	Male	15,189 (87.5)	61.7	30.16 (1)	< 0.001	0.042
	Female	2,056 (12.5)	55.5			
Media exposure	Yes	10,143 (58.5)	59.7	13.56 (1)	< 0.001	0.028
	No	7,102 (41.5)	62.5			
Wealth index	Poor	6,426 (37.7)	63.6	38.50 (2)	< 0.001	0.047
	Middle	3,436 (20.5)	61.0			
	Rich	7,383 (41.7)	58.4			
Digital inclusion	Low	9,758 (55.8)	68.4	569.96 (2)	< 0.001	0.182
	Medium	5,404 (31.8)	48.8			
	High	2,083 (12.3)	58.0			
Experience of child death	No	15,337 (89.2)	58.6	321.54 (1)	< 0.001	0.137

Characteristic	Category	n (%)	Limiting, %	χ^2 (df)	p	Cramér's V
Currently pregnant	Yes	1,908 (10.8)	80.0	179.72 (1)	< 0.001	0.102
	No	16,224 (94.1)	62.1			
Number of living children	Yes	1,021 (5.9)	41.0	3547.79 (2)	< 0.001	0.454
	0	1,707 (9.9)	2.0			
	1–3	13,703 (79.7)	63.4			
	≥ 4	1,835 (10.4)	97.9			
Contraceptive behavior/intent	Currently using	11,646 (67.0)	69.2	1074.23 (3)	< 0.001	0.250
	Nonuser, intends to use	2,063 (12.0)	40.6			
	Nonuser, no intention	3,206 (19.1)	47.8			
	Nonuser, unsure/other	330 (2.0)	28.8			
Husband's coresidence	Living with her	14,661 (84.1)	63.6	288.97 (1)	< 0.001	0.129
	Living elsewhere	2,584 (15.9)	46.3			
Age at marriage (years)	< 15	4,703 (27.9)	75.1	752.17 (2)	< 0.001	0.209
	15–17	6,754 (39.9)	60.7			
	≥ 18	5,788 (32.2)	48.8			
Ideal family size	0–2	12,951 (74.7)	60.6	1.94 (1)	0.164	0.011
	≥ 3	4,294 (25.3)	61.8			
Husband's education	No education	3,657 (21.5)	80.3	1010.48 (3)	< 0.001	0.242
	Primary	4,882 (28.6)	63.8			

Characteristic	Category	n (%)	Limiting, %	χ^2 (df)	p	Cramér's V
	Secondary	5,568 (32.7)	53.4			
	Higher	3,138 (17.2)	46.0			
Husband's occupation	Not working/other	671 (3.8)	67.1	521.65 (4)	< 0.001	0.174
	Agriculture	3,506 (19.3)	64.8			
	Manual labor	3,937 (22.2)	53.1			
	Business/services/sales	5,057 (30.3)	53.4			
	Professional/technical	4,074 (24.3)	73.2			
Respondent's working status	Not working	11,907 (67.8)	56.7	264.77 (1)	< 0.001	0.124
	Working	5,338 (32.2)	69.7			
Women's autonomy	Low	2,549 (14.0)	49.5			
	Medium	4,566 (26.2)	57.2			
	High	10,130 (59.7)	65.2			
Justifies wife-beating	No	14,990 (86.1)	60.2	20.66 (1)	< 0.001	0.035
	Yes	2,255 (13.9)	65.1			
Anxiety severity (GAD-7)	Minimal (0–4)	12,799 (74.5)	58.2			
	Mild (5–9)	3,751 (21.5)	68.2			
	Moderate/severe (10–21)	695 (4.0)	71.5			
Depression severity (PHQ-9)	Minimal (0–4)	12,354 (72.0)	58.9	76.69 (2)	< 0.001	0.067
	Mild (5–9)	4,067 (23.4)	65.4			

Characteristic	Category	n (%)	Limiting, %	χ^2 (df)	p	Cramér's V
	Moderate/severe (10–27)	824 (4.6)	68.8			

n, unweighted count; %, survey-weighted percentage. Limiting %, survey-weighted row percentage of women wanting no more children. Chi-square tests were computed on weighted contingency tables with weights normalized to sum to n, without continuity correction. Cramér's V, conventional thresholds: negligible < 0.10, weak 0.10–0.29, moderate 0.30–0.49, strong ≥ 0.50 . Overall, 10,566 (60.9%) women had a limiting preference, and 6,679 (39.1%) had a nonlimiting preference.

Feature selection

Boruta confirmed 27 of the 38 encoded features, left one tentative (husband not working or in another occupation) and rejected 10, comprising the seven division contrasts, religion, media exposure and attitudes toward wife beating (**Supplementary Table S2**). The age group presented 31.1% mean random forest importance, followed by having resided for 15 years or longer (19.4%), having living children (14.5%), having recently migrated (11.6%) and having a respondent's education level (6.4%). The VIFs among the 27 confirmed features ranged from 1.07 to 2.07, indicating no problematic collinearity (Additional file: Table S3). The 27 confirmed features were used for model training.

Model performance

All five models discriminated well on the test set, with AUCs ranging from 0.916 to 0.926 and overlapping confidence intervals (Table 2; Fig. 1). XGBoost had the highest cross-validated AUC (0.914 ± 0.009) and was selected as the final model; in the test set, it achieved an AUC of 0.926 (95% CI 0.917–0.934), a PR-AUC of 0.951, an accuracy of 84.2%, a precision of 89.4%, a sensitivity of 84.1% and a specificity of 84.3%. CatBoost performed almost identically (AUC 0.925), followed closely by logistic regression (0.919), random forest (0.918) and SVM (0.916). The held-out test AUCs were slightly higher than the corresponding cross-validated AUCs for all five models, although these differences were small. XGBoost correctly classified 1,778 of 2,113 women with a limiting preference and 1,126 of 1,336 without (**Supplementary Fig. S2**).

Table 2
Cross-validated and held-out test performance of the five classifiers for fertility-limiting preference

Model	CV AUC (mean $\pm$ SD)	Test AUC (95% CI)	PR-AUC	Accuracy	Precision	Sensitivity	Specificity	F1-score
XGBoost	0.914 $\pm$ 0.009	0.926 (0.917–0.934)	0.951	84.2	89.4	84.1	84.3	0.867
CatBoost	0.913 $\pm$ 0.009	0.925 (0.917–0.934)	0.951	84.1	89.1	84.3	83.7	0.866
Random Forest	0.907 $\pm$ 0.008	0.918 (0.909–0.928)	0.945	83.6	88.2	84.5	82.2	0.863
SVM	0.907 $\pm$ 0.010	0.916 (0.907–0.926)	0.945	83.8	88.5	84.6	82.6	0.865
Logistic Regression	0.907 $\pm$ 0.010	0.919 (0.910–0.929)	0.947	83.5	88.9	83.5	83.5	0.861

CV AUC, mean $\pm$ SD of the area under the receiver operating characteristic curve across 10 stratified cross-validation folds on the training set ($n = 13,796$), the model selection criterion. Test metrics were computed once on the held-out test set ($n = 3,449$) for the limiting class; the 95% CI is from 1,000 bootstrap resamples. PR-AUC, area under the precision–recall curve. The models are ordered by the CV AUC.

Model interpretation

Fifteen features accounted for 91.4% of the total mean absolute SHAP in the XGBoost model (Fig. 2; Additional file: Table S4 and Fig. S4). The age group (24.3%) and number of living children (13.1%) dominated, followed by nonuse of contraception with no intention to use (7.5%), residence of 15 years or longer (7.1%), ideal family size (6.6%), age at marriage (6.3%) and household size (6.1%). The respondent's education (3.7%), recent migration (3.3%), current pregnancy (3.1%), intention to begin contraception (3.0%), husband's education (2.2%), male household headship (2.0%), digital inclusion (1.6%) and anxiety severity (1.5%) contributed the remainder; depression severity (0.4%) and child loss (0.3%) contributed the least.

The beeswarm plot (Fig. 2) shows the direction of these contributions. Older age and higher parity produced the largest positive contributions to the log-odds of limiting, and childlessness had the largest negative contributions. A residence of 15 years or longer increased the predicted probability of limiting, whereas recent migration decreased it. A larger ideal family size, later marriage, higher education of the woman and her husband, and nonuse of contraception, whether with an intention to begin or with no intention, shifted predictions toward wanting another child. A small household size also reduced the predicted probability of limiting. Current pregnancy, which was associated with a lower proportion of limitations in bivariate analysis, increased the predicted probability of being limited once age and parity were accounted for. Female

household headship and greater anxiety severity made small positive contributions. Local explanations for the women with the highest (0.996) and lowest (0.001) predicted probabilities are shown in **Supplementary Fig. S3**.

Discussion

This study aimed to predict fertility-limiting preferences among currently married women in Bangladesh and to identify which characteristics contributed most strongly to the model's predictions. Using 17,245 women from the 2022 BDHS, three findings answer the questions posed. First, fertility-limiting preference was highly predictable: all five algorithms achieved test AUCs above 0.91, and XGBoost reached 0.926, with balanced sensitivity and specificity. Second, the predictive signal was concentrated in a compact life-course core of age, parity, contraceptive behavior, residential duration, ideal family size, marriage timing and household size, which together accounted for approximately 70% of the model-attributed importance. Third, SHAP provided directionally explicit explanations of the model predictions and showed that several characteristics associated with fertility preference in bivariate analyses, including digital inclusion and mental health symptoms, made relatively small contributions to the final model after the other measured characteristics were incorporated.

In addition to these findings, this study adds to the small literature that applies ML to fertility preferences in three ways. To our knowledge, no previous study has applied an explainable ML framework to fertility preferences in Bangladesh or in another late-transition South Asian population, where the question is who wants to stop rather than why fertility is high; previous DHS-based applications in Nigeria and Somalia [16, 17] addressed high-fertility settings. It tests whether digital inclusion and screening for anxiety and depression contribute additional predictive information beyond the other measured characteristics, which has not been examined in a fertility-preference prediction model. This finding shows that confining feature selection and encoding to the training set and using class weighting instead of synthetic oversampling yields discrimination comparable to that of studies without these safeguards, so the reported performance is unlikely to be an artifact of the evaluation design.

The discrimination achieved is comparable to that reported from high-fertility settings, where Sani et al. obtained an AUC of 0.89 [17] and Tadese et al. an AUROC of 0.92 [16], but was reached without synthetic oversampling and with feature selection confined to the training set, so the present estimates are less likely to be optimistic. Comparable discrimination in a late-transition population is consistent with intentions here being strongly structured by life-course position, which keeps the classification problem learnable even though pronatalist norms no longer hold preferences high across the life course. The near-equivalence of logistic regression (AUC 0.919) and gradient boosting (0.926) suggests that the dominant relationships are largely monotone, with the ensembles capturing a modest additional nonlinear increment, a pattern consistent with systematic review evidence that machine learning offers no general performance benefit over logistic regression for clinical prediction [15, 38].

The dominance of age and parity is consistent with previous studies of fertility preference [2, 16, 17]: the desire to stop childbearing increases steeply as women approach or exceed their desired family size. The education gradients for both spouses mirror the role of schooling in reducing the desired family size and

increasing receptivity to family planning documented across sub-Saharan Africa [4, 39, 40]. The contribution of ideal family size is notable because it showed no bivariate association: its predictive value emerged only conditional on parity, which is consistent with preferences reflecting the gap between achieved and desired family size rather than either alone [41]. Early marriage operates through exposure: women who are married before age 15, who make up 27.9% of the sample, reach any given age with higher parity. The fact that current contraceptive users are disproportionately limiters, whereas nonusers intending to begin are disproportionately spacers or starters, indicates coherence between intention and practice; the negative contribution for nonusers with no intention likely reflects that this group includes women who are pregnant, postpartum or seeking pregnancy.

Internal migration was among the strongest predictors in both Boruta and SHAP, although its SHAP contribution (approximately 11% across the three contrasts) was smaller than its random forest importance suggested. Recent migrants were overwhelmingly nonlimiting (76.4%) and long-term residents overwhelmingly limiting (89.5%), and SHAP showed opposite-signed contributions for recent migration and long-term residence within the fitted model, alongside the effects of age and parity. This pattern is interpretable through selection and life-course mechanisms [21]: much internal migration in Bangladesh is marriage-related or rural-to-urban labor migration that selects young couples at the start of family building, whereas long tenure plausibly indexes completed family formation.

The current pregnancy lowered the crude proportion of limits but increased the predicted probability once age and parity were held constant, which may reflect differences in the relationship between pregnancy, age, parity and stated fertility preference; however, the cross-sectional design does not allow the mechanism to be established. The digital inclusion and mental health screening scores, which were examined here for the first time in a fertility-preference prediction model, were associated with preference in bivariate analysis but contributed little independent signal, in contrast to the associations reported in the wider literature [22, 24], suggesting that their crude associations here largely reflect the age, education and parity profile of digitally included or symptomatic women rather than independent pathways. Similarly, child loss was associated with limiting preference in bivariate analysis (80.0% vs 58.6%) but ranked last in SHAP, indicating that it added relatively little predictive information beyond the other variables included in the model [23].

With 60.9% of married women wanting no more children but 19.1% being nonusers with no intention to use, the model's high-precision identification of limiters (89.4%) could support prioritization of counseling on long-acting and permanent methods, whereas recent-migrant, low-parity women are candidates for spacing services. The relatively small contribution of division, wealth and residence suggests that the challenge may be less one of geographic or economic differences in the desire to limit than of converting later-life limiting preferences into continuous contraceptive practices [8, 10]. Predictions of this kind are associations learned from cross-sectional data and should not be read as causal effects [38].

Strengths and limitations

Strengths include a large nationally representative sample; a predictor set spanning migration; digital, relational and mental health domains; a leakage-safe pipeline with model selection on cross-validated performance; and transparent SHAP interpretation [30, 36]. Several limitations should be noted. The cross-

sectional design precludes causal inference, and predictors such as contraceptive intention are plausibly codetermined with the outcome. Fertility preferences are self-reported at a single time point and may be revised over the life course. The sample was restricted to women receiving long questionnaires, which were allocated to two-thirds of households randomly; thus, representativeness should be preserved. However, women undecided about a further child were combined with those wanting more, and infecund women were excluded. The migration measure captures the duration but not the origin, reason or distance of moves. Feature selection preceded cross-validation, so the cross-validated AUCs may be mildly optimistic; however, the held-out test set remained untouched during feature selection and model tuning. The models were trained on unweighted observations, as is usual for prediction, and have not been externally validated on other surveys [42].

Conclusions

This study found that fertility-limiting preferences among currently married women in Bangladesh can be predicted with good discrimination from routinely collected survey characteristics. The XGBoost model achieved a test AUC of 0.926, while age, number of living children, contraceptive behavior, duration of residence, ideal family size, age at marriage and household size accounted for most of the predictive signal. Digital inclusion, mental health symptoms, wealth and geographic variables contributed relatively little to the fitted model after the other characteristics were considered.

Among the predictors, duration of residence and contraceptive intention may be useful for identifying groups with differing fertility preferences and informing counseling needs assessment. However, these findings represent predictive associations rather than causal effects, and external validation in subsequent BDHS rounds and other populations is needed before such models are considered for individual-level service delivery.

Abbreviations

AUC

area under the receiver operating characteristic curve

BDHS

Bangladesh Demographic and Health Survey

CI

confidence interval

DHS

demographic and health survey

GAD-7

generalized anxiety disorder 7-item scale

ML

machine learning

NIPORT

National Institute of Population Research and Training

PHQ-9

Patient Health Questionnaire 9-item scale

PR-AUC

area under the precision–recall curve

SHAP

Shapley additive explanations

SVM

support vector machine

TRIPOD + AI

transparent reporting of a multivariable prediction model for individual prognosis or diagnosis + artificial intelligence

VIF

variance inflation factor

XGBoost

extreme gradient boosting.

Declarations

Conflicts of Interest:

The author(s) declare no known competing interests that could have appeared to influence the work reported in this paper.

Ethical approval and consent to participate:

This study analyzed fully anonymized, publicly available secondary data from the 2022 BDHS, obtained with permission from the DHS Program (<https://dhsprogram.com>). The original survey protocol was reviewed and approved by the ICF Institutional Review Board and the Bangladesh Medical Research Council, and all procedures were carried out in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants before the interview. No additional ethical approval was required for this secondary analysis of de-identified data.

Funding statement:

No funding or incentives were received for this research project and publication of the manuscript.

Author Contribution

Prayas Sur Antu (P.S.A.) and Md. Mahbub-Ur Rashid (M.M.R.) contributed to conceptualization, methodology, data curation, formal analysis, validation, visualization, and writing – original draft. Ujjoyene Talapatra Athai

(U.T.A.) contributed to writing – original draft. Md. Hazrat Ali (M.H.A.) contributed to supervision, methodology, and writing – review & editing. All authors read and approved the final manuscript.

Acknowledgement

The authors express gratitude to the Bangladesh Demographic and Health Survey (BDHS) Program and all survey participants for their invaluable contributions to this research.

Data availability statement:

The data that support the findings of this study are openly available in the Bangladesh Demographic and Health Survey (BDHS) 2022 at <https://dhsprogram.com/>.

References

1. Cleland J, Machiyama K, Casterline JB. Fertility preferences and subsequent childbearing in Africa and Asia: a synthesis of evidence from longitudinal studies in 28 populations. *Popul Stud (Camb)*. 2020;74(1):1–21. <https://doi.org/10.1080/00324728.2019.1672880>.
2. Ahinkorah BO, Seidu A-A, Armah-Ansah EK, Ameyaw EK, Budu E, Yaya S. Socioeconomic and demographic factors associated with fertility preferences among women of reproductive age in Ghana: evidence from the 2014 Demographic and Health Survey. *Reprod Health*. 2021;18:2. <https://doi.org/10.1186/s12978-020-01057-9>.
3. Muluneh MW, Moyehodie YA. Determinants of desire for more children among women in Ethiopia. *BMC Womens Health*. 2021;21:408. <https://doi.org/10.1186/s12905-021-01563-3>.
4. Bongaarts J. Trends in fertility and fertility preferences in sub-Saharan Africa: the roles of education and family planning programs. *Genus*. 2020;76:32. <https://doi.org/10.1186/s41118-020-00098-z>.
5. Odusina EK, Ayotunde T, Kunnuji M, Ononokpono DN, Bishwajit G, Yaya S. Fertility preferences among couples in Nigeria: a cross sectional study. *Reprod Health*. 2020;17:92. <https://doi.org/10.1186/s12978-020-00940-9>.
6. Bearak J, Popinchalk A, Ganatra B, Moller A-B, Tunçalp Ö, Beavin C, et al. Unintended pregnancy and abortion by income, region, and the legal status of abortion: estimates from a comprehensive model for 1990–2019. *Lancet Glob Health*. 2020;8(9):e1152–61. [https://doi.org/10.1016/S2214-109X\(20\)30315-6](https://doi.org/10.1016/S2214-109X(20)30315-6).
7. Duminy J, Cleland J, Harpham T, Montgomery MR, Parnell S, Speizer IS. Urban family planning in low- and middle-income countries: a critical scoping review. *Front Glob Womens Health*. 2021;2:749636. <https://doi.org/10.3389/fgwh.2021.749636>.
8. Hellwig F, Moreira LR, Silveira MF, Vieira CS, Rios-Quitizaca PB, Masabanda M, et al. Policies for expanding family planning coverage: lessons from five successful countries. *Front Public Health*. 2024;12:1339725. <https://doi.org/10.3389/fpubh.2024.1339725>.

9. National Institute of Population Research and Training (NIPORT), ICF. Bangladesh Demographic and Health Survey 2022: final report (FR386). Dhaka, Bangladesh, and Rockville, Maryland, USA: NIPORT and ICF. 2024. <https://www.dhsprogram.com/pubs/pdf/FR386/FR386.pdf>. Accessed 15 Jun 2024.
10. Streatfield PK, Kamal N. Population and family planning in Bangladesh. *J Pak Med Assoc.* 2013;63(4 Suppl 3):S73-81. PMID: 24386734.
11. Bairagi R. Effects of sex preference on contraceptive use, abortion and fertility in MATLAB, Bangladesh. *Int Fam Plan Perspect.* 2001;27(3):137–43. <https://doi.org/10.2307/2673835>.
12. Calhoun LM, Nanda P, Speizer IS, Jain M. The effect of family sex composition on fertility desires and family planning behaviors in urban Uttar Pradesh, India. *Reprod Health.* 2013;10:48. <https://doi.org/10.1186/1742-4755-10-48>.
13. Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Med (Kaunas).* 2020;56(9):455. <https://doi.org/10.3390/medicina56090455>.
14. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inf Decis Mak.* 2019;19:281. <https://doi.org/10.1186/s12911-019-1004-8>.
15. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12–22. <https://doi.org/10.1016/j.jclinepi.2019.02.004>.
16. Tadese ZB, Nimani TD, Mare KU, Gubena F, Wali IG, Sani J. Exploring machine learning algorithms for predicting fertility preferences among reproductive age women in Nigeria. *Front Digit Health.* 2025;6:1495382. <https://doi.org/10.3389/fdgth.2024.1495382>.
17. Sani J, Halane S, Ahmed AM, Ahmed MM. Application of machine learning algorithms and SHAP explanations to predict fertility preference among reproductive women in Somalia. *Sci Rep.* 2025;15:26301. <https://doi.org/10.1038/s41598-025-04704-y>.
18. Kassaw EA, Abate BB, Enyew BM, Sendekie AK. The application of machine learning approaches to classify and predict fertility rate in Ethiopia. *Sci Rep.* 2025;15:2562. <https://doi.org/10.1038/s41598-025-85695-8>.
19. Moulaei K, Mahboubi M, Ghorbani Kalkhajeh S, Kazemi-Arpanahi H. Prediction of childbearing tendency in women on the verge of marriage using machine learning techniques. *Sci Rep.* 2024;14:20811. <https://doi.org/10.1038/s41598-024-71854-w>.
20. Mia M, Akter S, Sraboni TN, Islam MM, Tawabunnahar M, Maniruzamman M, et al. Predicting short birth intervals in Bangladesh using stacked machine learning and SHAP explainability: evidence from BDHS 2022. *Reprod Health.* 2026;23:154. <https://doi.org/10.1186/s12978-026-02367-0>.
21. Kulu H. Migration and fertility: competing hypotheses re-examined. *Eur J Popul.* 2005;21(1):51–87. <https://doi.org/10.1007/s10680-005-3581-8>.
22. Rotondi V, Kashyap R, Pesando LM, Spinelli S, Billari FC. Leveraging mobile phones to attain sustainable development. *Proc Natl Acad Sci U S A.* 2020;117(24):13413–20. <https://doi.org/10.1073/pnas.1909326117>.

23. LeGrand T, Koppenhaver T, Mondain N, Randall S. Reassessing the insurance effect: a qualitative analysis of fertility behavior in Senegal and Zimbabwe. *Popul Dev Rev.* 2003;29(3):375–403. <https://doi.org/10.1111/j.1728-4457.2003.00375.x>.
24. Carlsson E, Kim W. Does depression matter for fertility intentions and their realization? An analysis of Norwegian and Swedish men and women in coresidential unions. *Genus.* 2026;82:3. <https://doi.org/10.1186/s41118-026-00283-6>.
25. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092–7. <https://doi.org/10.1001/archinte.166.10.1092>.
26. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606–13. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>.
27. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw.* 2010;36(11):1–13. <https://doi.org/10.18637/jss.v036.i11>.
28. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Guyon I, Von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al. editors. *Advances in Neural Information Processing Systems 30*. Red Hook, NY: Curran Associates; 2017. pp. 4765–74.
29. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378. <https://doi.org/10.1136/bmj-2023-078378>.
30. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health.* 2019;22(7):808–15. <https://doi.org/10.1016/j.jval.2019.02.012>.
31. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20(3):273–97. <https://doi.org/10.1007/BF00994018>.
32. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32. <https://doi.org/10.1023/A:1010933404324>.
33. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM; 2016. pp. 785 – 94. <https://doi.org/10.1145/2939672.2939785>
34. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. In: Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R, editors. *Advances in Neural Information Processing Systems 31*. Red Hook, NY: Curran Associates; 2018. pp. 6639–49.
35. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell.* 2020;2(1):56–67. <https://doi.org/10.1038/s42256-019-0138-9>.
36. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. *Clin Transl Sci.* 2024;17(11):e70056. <https://doi.org/10.1111/cts.70056>.
37. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825–30.

38. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: making sense of the differences. *Knee Surg Sports Traumatol Arthrosc.* 2022;30(3):753–7. <https://doi.org/10.1007/s00167-022-06896-6>.
39. Berger H, Dasré A. Religious affiliation, education, and fertility in sub-Saharan Africa. *World Dev.* 2024;184:106723. <https://doi.org/10.1016/j.worlddev.2024.106723>.
40. Götmark F, Andersson M. Human fertility in relation to education, economy, religion, contraception, and family planning programs. *BMC Public Health.* 2020;20:265. <https://doi.org/10.1186/s12889-020-8331-7>.
41. Nitsche N, Hayford SR. Preferences, partners, and parenthood: linking early fertility desires, marriage timing, and achieved fertility. *Demography.* 2020;57(6):1975–2001. <https://doi.org/10.1007/s13524-020-00927-y>.
42. Bozkurt S, Cahan EM, Seneviratne MG, Sun R, Lossio-Ventura JA, Ioannidis JPA, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. *J Am Med Inf Assoc.* 2020;27(12):1878–84. <https://doi.org/10.1093/jamia/ocaa164>.

Figures

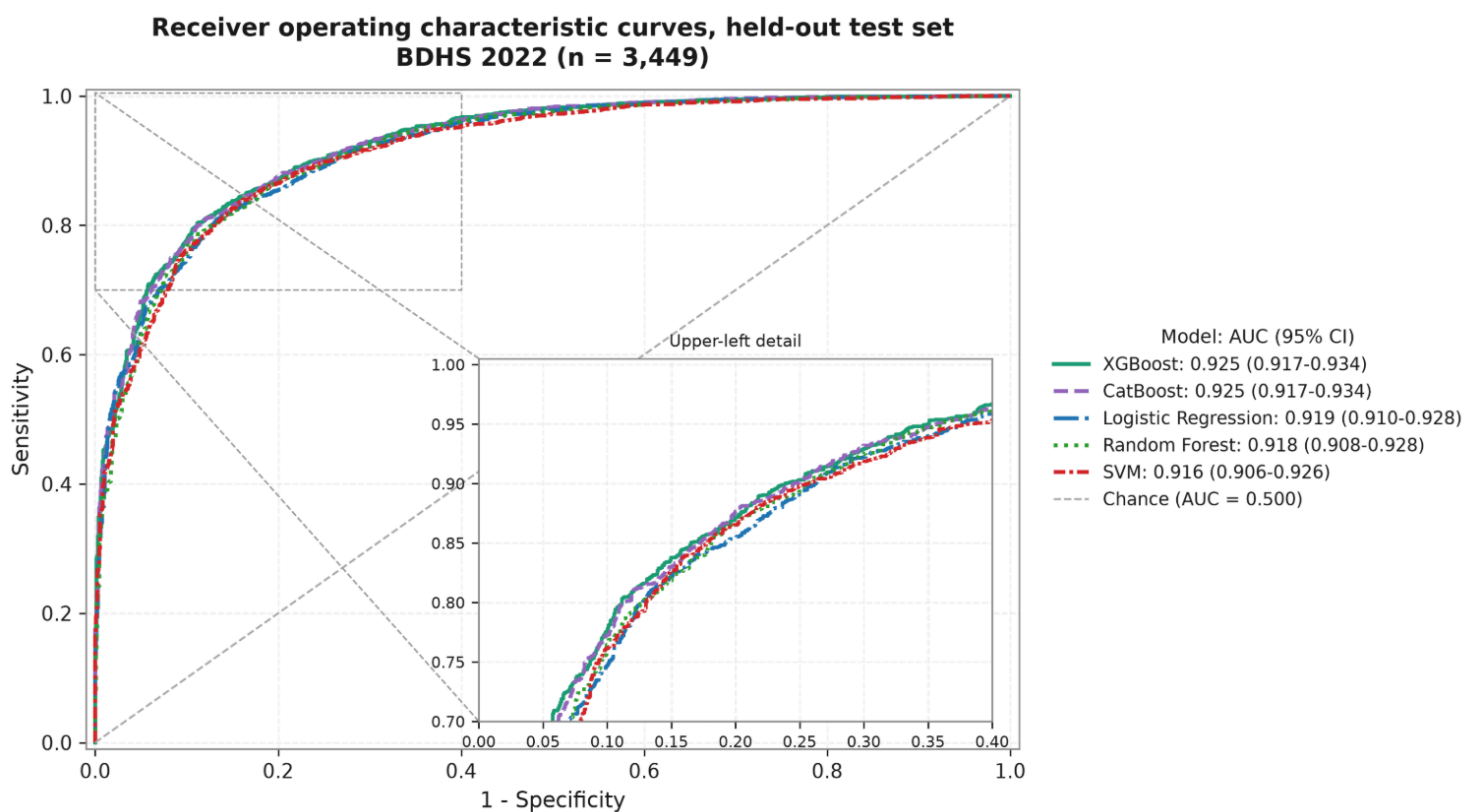


Figure 1

Receiver operating characteristic curves of the five classifiers on the held-out test set. Curves are shown for the test set (n = 3,449). The legend shows the area under the curve (AUC) with a 95% bootstrap

confidence interval (1,000 resamples) for each model; the inset shows the upper-left region. The dashed diagonal represents chance (AUC 0.500).

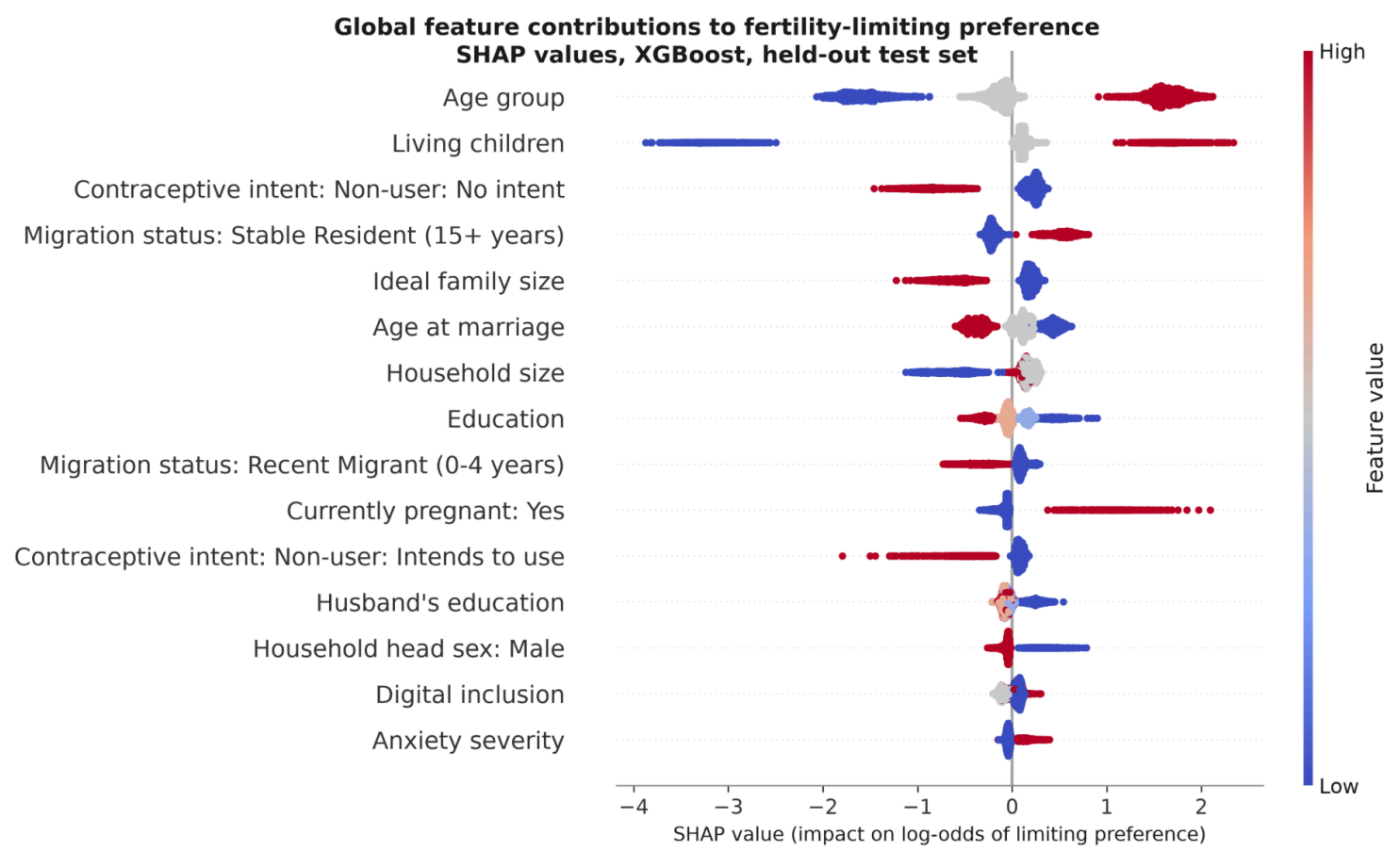


Figure 2

SHAP summary plot of the final XGBoost model on the held-out test set. Each point represents one woman in the test set (n = 3,449). The features are ordered by the mean absolute SHAP value and restricted to the 15 features jointly accounting for 91.4% of the total. The horizontal position is the contribution to the log-odds of a limiting preference; color indicates the feature value (blue low, red high; for binary contrasts, red = category present). The full rankings are given in Additional file: Table S4.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFertilityPreferences.docx](#)