scientific reports

OPEN

Employing machine learning models to predict pregnancy termination among adolescent and young women aged 15–24 years in East Africa

Getanew Aschalew Tesfa[1]✉, Abel Desalegn Demeke[2], Binyam Tariku Seboka[1], Tsion Mulat Tebeje[1], Mulugeta Desalegn Kasaye[3], Behailu Taye Gebremeskele[4], Samuel Hailegebreal[5] & Habtamu Setegn Ngusie[6]

Pregnancy termination is still a sensitive and continuing public health issue due to several political, economic, religious, and social concerns. This study assesses the applications of machine learning models in the prediction of pregnancy termination using data from eleven national datasets in East Africa. Nine machine learning models, namely: Random Forests (RF), Decision Tree, Logistic Regression, Support Vector Machine, eXtreme Gradient Boosting (XGB), AdaBoost, CatBoost, K-nearest neighbor, and feedforward neural network models were used to predict pregnancy termination, with six evaluation criteria utilized to compare their performance. The pooled prevalence of pregnancy termination in East Africa was found to be 4.56%. All machine learning models had an accuracy of at least 71.8% on average. The RF model provided accuracy, specificity, precision, and AUC of 92.9%, 0.87, 0.91, and 0.93, respectively. The most important variables for predicting pregnancy termination were marital status, age, parity, country of residence, age at first sexual activity, exposure to mass media, and educational attainment. These findings underscore the need for a tailored approach that considers socioeconomic and regional disparities in designing policy initiatives aimed at reducing the rate of pregnancy terminations among younger women in the region.

Keywords Abortion, Artificial intelligence, Pregnancy termination, Women

Abbreviations

AI	Artificial intelligence
AUC	Area under the curve
DHS	Demographic and health survey
DT	Decision tree
ML	Machine learning
RF	Random forest
ROC	Receiver operating characteristic
SVM	Support vector machine

The term pregnancy termination is often used interchangeably with abortion[1]. Induced abortion is the deliberate surgical or medical termination of a viable fetus, while spontaneous abortions, also known as miscarriages, happen when a fetus or an embryo is lost because of external circumstances[2]. Abortion can occur spontaneously

[1]School of Public Health, College of Medicine and Health Science, Dilla University, Dīla, Ethiopia. [2]Nursing department, college of Medicine and Health Science, Dilla University, Dīla, Ethiopia. [3]Department of Health Informatics, College of Medicine and Health Science, Wollo University, Dessie, Ethiopia. [4]Department of Medical Laboratory Science, College of Medicine and Health Science, Dilla University, Dīla, Ethiopia. [5]Department of Health Informatics, College of Medicine and Health Science, Wachamo University, Hosaina, Ethiopia. [6]School of Public Health, College of Medicine and Health Science, Woldia University, Woldia, Ethiopia. ✉email: getanewaschalew@gmail.com

or intentionally and it can be safe or risky. Risky abortions, in particular, can lead to complications such as hemorrhage, infection, and uterine perforation, which can have significant health consequences[3,4].

According to the United Nations Population Fund 2022 report, nearly half of all pregnancies between 2015 and 2019 were unintended. This corresponds to approximately 121 million unintended pregnancies annually. More than 60% of unintended pregnancies result in abortion, according to the report. Furthermore, 45% of abortions carried out worldwide are unsafe[5].

Globally, unsafe pregnancy termination is responsible for 4.7–13.2% of annual maternal deaths[6]. In developed regions, 30 women are estimated to die for every 100,000 unsafe abortions, while the number rises to 220 deaths per 100,000 unsafe abortions in developing countries[7]. In developing nations, unsafe abortion procedures are the leading contributor to maternal mortality and morbidity[8]. Abortion is among the top five causes of maternal death in Low and Middle-Income Countries (LMICs). It is linked to pregnancy and childbirth-related complications[9].

Pregnancy termination is never an easy decision, particularly in LMICs, where women's decisions may be influenced by several factors such as restrictive abortion legislation, patriarchal communities, religious considerations, cultural mores, and economic circumstances[10]. The legal status of abortion varies widely between nations; most permit it under specific conditions, while some completely prohibit it. Most developed nations permit the practice. Countries set restrictions on abortion, often allowing it only in particular circumstances such as financial difficulties, risks to the woman's physical or mental health, or the presence of fetal abnormalities[11–13].

Nowadays, researchers and medical experts are paying attention to artificial intelligence (AI) in the healthcare industry[14]. Applications of AI in healthcare focus on several areas, including clinical decision-making, patient data and diagnostics, predictive medicine, health services management, and much more[15–19].

One of the most interesting applications of AI is machine learning (ML), and many businesses are trying to use it. Machine learning is growing in popularity. It can be applied in a variety of contexts, including business and healthcare, and leverages algorithms to support data-driven learning[20]. Machine learning is a state-of-the-art method for analyzing vast databases, discovering pertinent data, and developing predictive, clustering, and association models[21,22]. Recently machine learning is a popular computing technique that is also being applied in public health and medicine[22–24].

Various studies are conducted in different nations to identify the determinants of pregnancy termination[25–30]. However, they fail to report information that is unknown in advance. Employing machine learning approaches encompasses a wider range of methods beyond just the commonly used statistical techniques such as logistic regression modeling.

To our knowledge, there is limited evidence about predicting pregnancy termination among young girls utilizing machine learning methods. Therefore, this study aimed to investigate pregnancy termination by employing various machine learning techniques among adolescent and young women aged 15–24 years old in East African countries based on the most recent Demographic and Health Surveys. Our study's findings could be useful in guiding focused public health initiatives for improving maternal health and reducing the global maternal mortality rate. In this study, we employed strong machine learning techniques to develop a prediction model using data from extensive demographic and health surveys of eleven East African countries. Additionally, we also utilized feature importance selection to identify the most important factors that are linked to the prediction of pregnancy termination.

Methods and materials

Data source

The data was obtained from the Measure Demographic and Health Surveys (DHS) official database that was collected in eleven East African countries (Burundi, Ethiopia, Kenya, Comoros, Rwanda, Tanzania, Madagascar, Malawi, Zambia, Zimbabwe, and Uganda). DHS collected the data in each nation, and we examined the most recent available data of each nation, which were collected from 2012 to 2022. The data is available from the DHS's official website (https://dhsprogram.com/data/available-datasets.cfm). Since the data set is freely available from the measure DHS official database, anyone can access it through reasonable request.

Every five years, the DHS surveys are conducted using standardized, pretested, and validated questionnaires in low- and middle-income nations (over 90 countries). The DHS surveys adhere to the same standard approach for sampling, questionnaires, data collection, and coding. A stratified two-stage cluster sampling technique is used in the survey. This study focused on adolescent and young women (aged 15–24 years old) in East Africa (Table 1).

Outcome variable

In this study, the outcome variable "pregnancy termination" was measured as a binary outcome, which was derived from the DHS question "Have you ever had a terminated pregnancy?". Thus, pregnancy termination was measured as yes (coded as 1) or no (coded as 0) for all the models. In DHS, pregnancy termination includes both spontaneous and induced abortions.

Independent variables

The predictor variables (features) considered in this study are participants age, educational status, age at first sex, marital status, distance to nearest health facility, country of residence, place of residence, mass media exposure, community level educational level, total number of children ever born, number of children born in the last five years, wealth index, community level poverty, and community level mass media exposure status (Table 2).

Countries	Unweighted sample		Weighted sample	
	Frequency	Percent	Frequency	Percent
Burundi	7218	9.60	7103	9.51
Ethiopia	6401	8.51	6143	8.23
Kenya	12,166	16.18	12,026	16.11
Comoros	2278	3.03	2307	3.09
Madagascar	7881	10.48	7906	10.59
Malawi	10,367	13.78	10,422	13.96
Rwanda	5252	6.98	5225	7.00
Tanzania	5852	7.78	5810	7.78
Uganda	8058	10.71	8086	10.83
Zambia	5799	7.71	5733	7.68
Zimbabwe	3938	5.24	3895	5.22
Total	75,210	100	74,656	100

Table 1. Total study sample per country.

Predictor variables/features	Measures
Participants age	15–19, 20–24
Age at first sex	<15, ≥ 15
Educational level	No formal education, primary, secondary and above
Religion	Christian, Muslim, others
Working status	Not working, working
Wealth index	Poor, middle, rich
Smoking status	No, yes
Place of residence	Urban, rural
Distance to nearest HF	Not a big problem, big problem
Marital status	Unmarried, married, divorced/widowed/separated
Mass media exposure	Not exposed, exposed
Parity	No birth, 1–4, ≥ 5
Children born in the last 5 years	No birth, one, ≥ 2
Community-level educational status	Low, high
Community level poverty	High, low
Community-level mass media exposure	Low, high

Table 2. Description of independent variables.

Data preprocessing

Data pre-processing, which entails altering or encoding the data to make it appropriate for computer interpretation, is the first step in the machine-learning process. Every categorical and string variable was converted to a number value. Multiple imputation technique was used to handle missing values. The next stage after creating the final dataset was to pre-process the data using a variety of techniques. This process involved selecting the best model, fine-tuning its parameters, and creating relevant features. We implemented a continual improvement method for our models in our machine learning workflow. We improved our models through an iterative method. The main stages of our workflow are shown in Fig. 1.

Feature selection

Variable selection is a crucial component that helps in building prediction models free from biases, correlated variables, and unwanted noise[31]. We assessed various feature selection techniques and discovered that the Boruta algorithm worked effectively. The Boruta algorithm is a wrapper method built around the random forest classification algorithm. It attempts to include all the important, interesting features found in the dataset concerning an outcome variable[32]. Additionally, this feature selection algorithm is relevant, and able to identify both highly and weakly relevant variable features from the dataset[33]. Therefore, Feature selection was employed in R version 4.3.2 software with the R Boruta package.

Model building

Data balancing

In data mining and machine learning, class imbalance presents a problem that reduces the classification accuracy of instances belonging to minority classes[34]. Balancing the dataset makes it easier to train a model since

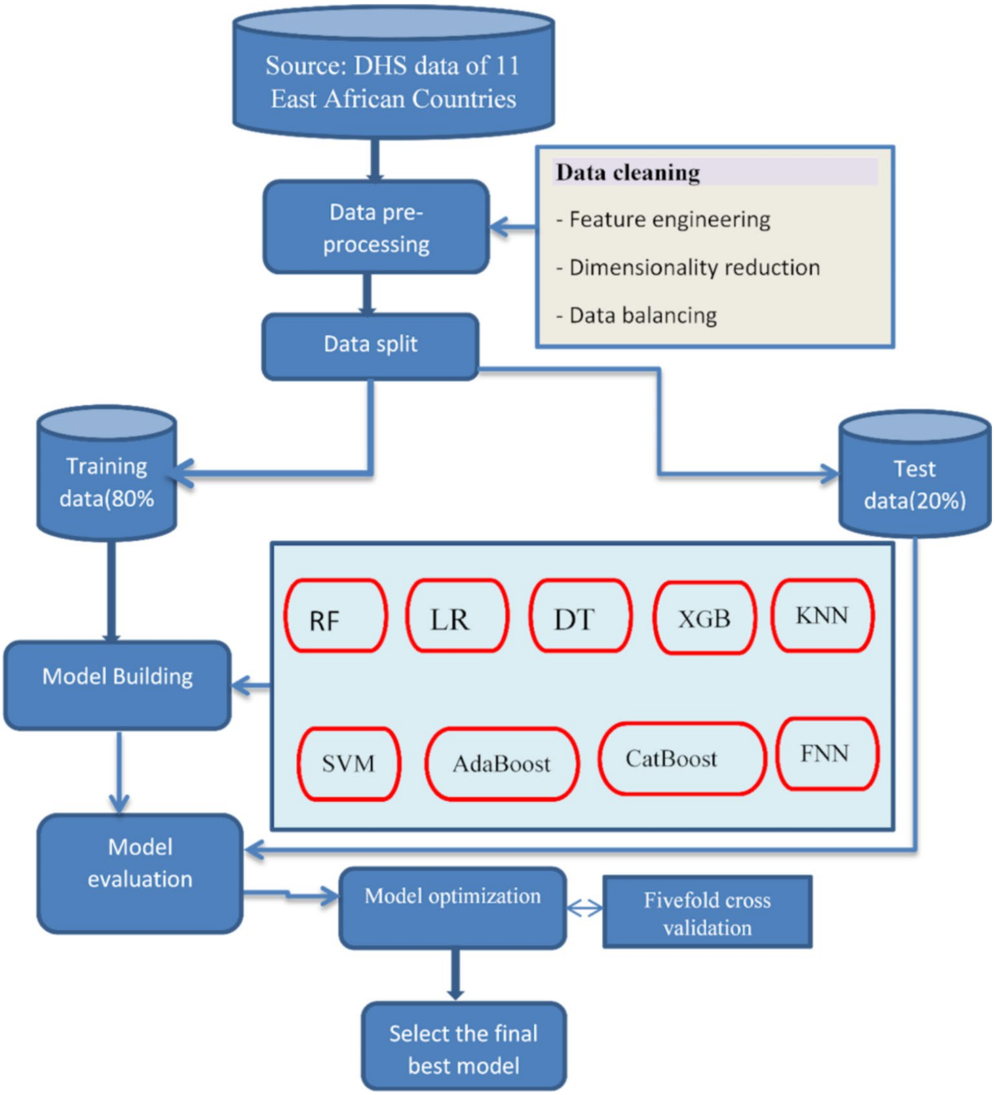

Fig. 1. The main steps of the workflow.

it prevents the model from getting biased toward one class. In other words, just because the model has more data, it won't automatically favor the majority class[35,36]. Therefore, we employed five data balancing techniques such as under sampling, over-sampling, near miss algorithm, adaptive synthetic sampling (ADASYN), and BorderLine synthetic Minority over-sampling Technique (BL-SMOTE) to improve the predictive model's performance and lessen the class imbalance. Consequently, a comprehensive evaluation of model performance and comparisons of all machine learning models were carried out. Generally, both the balanced and imbalanced data are shown in Fig. 2.

Next, the grid search optimization technique was utilized to optimize the hyperparameters. The objective of hyperparameter optimization is to refine and improve the model that delivers the highest and most accurate performance on a validation set.

Data splitting and model development

The dataset was split up into two groups, training and test groups were created by randomly assigning the dataset to each group. 80% of the entire dataset was made up of the training group. The training group was used as a basis for developing the prediction model. And 20% of the whole dataset was put aside as the test group to evaluate the model. Then, we employed a variety of appropriate machine learning algorithms, including Random Forests (RF), Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGB), AdaBoost, CatBoost, K-nearest neighbor (KNN), and FNN using python 3.12.2, implemented with Jupyter lab.

To demonstrate how well the models function in terms of predicting the pregnancy termination status, metrics for model accuracy such as sensitivity, specificity, accuracy, area under the receiver operating characteristic curve (AUC), and other learning metrics (F1 score and precision) were used since our target variable has a binary

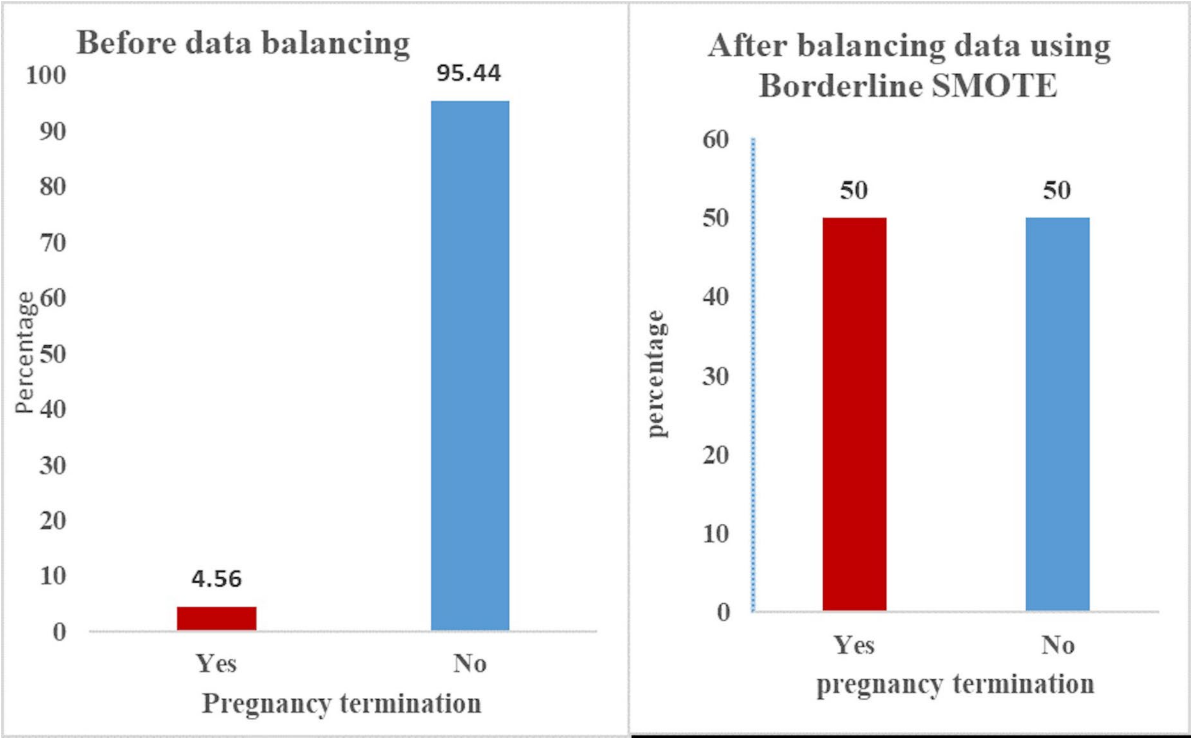

Fig. 2. Data comparison before and after applying the data balancing technique.

outcome. The final model was the best-performing model determined by combining metrics for performance evaluation.

Sensitivity or recall is the proportion of subjects expected to be positive among all those who are truly positive, whereas specificity is the proportion of subjects predicted to be negative among all those who are truly negative. And precision is defined as the proportion of subjects who are actually positive out of all those who are predicted to be positive[37].

The area under the curve (AUC) of the receiver operating characteristic (ROC) was used to assess how well the model could differentiate between instances. ROC curves plot sensitivity versus specificity across a range of thresholds to evaluate the capacity to predict a binary result. Specifically, the ROC approach involves representing the value of the sensitivity as a function of (1-specificity) for all possible threshold values and joining the points with a curve. The ROC curve provides the true positive rate as a function of the false positive rate for the same group. The classifier is better at differentiating between positive and negative classes if its curve is more angled toward the upper left corner. The AUC represents the classifier's capacity to distinguish between classes and is used as a summary of the ROC curve[38,39]. AUC values range from 0 to 1. A prediction is considered no better than chance if it is less than 0.5, poor if between 0.5 and 0.7, acceptable between 0.70 and 0.79, excellent between 0.8 and 0.89, and outstanding if greater than 0.9[40].

Model interpretability

After constructing and evaluating the predictive models, independent predictors (features) are used to predict pregnancy termination among adolescent and young women. To enhance the accuracy and relevance of these predictions, the model that demonstrated the best performance was utilized to identify and select the most significant features influencing the likelihood of pregnancy termination. This process involved analyzing various predictors to determine which ones had the greatest impact, thereby refining the model's ability to provide precise predictions. Features are arranged in descending order of feature importance. The feature that has the greatest feature importance score is the most significant and shows how important it is globally to predict pregnancy termination.

Additionally, SHAP (SHapley Additive exPlanations) analysis was used to show how each attribute affected the predictive model. SHAP analysis is a useful technique for figuring out the importance and input of specific features in a machine learning.

Software

STATA version 15 was used to recode the variables; Python 3.12 with Jupyter lab was used for further data management, to apply machine learning classifiers on the dataset, and to measure the evaluation criteria. The most important predictor variables were then identified using RStudio based on the model that performed the best. Additionally, feature selection was done by RStudio using the Boruta algorithm.

Results

Characteristics of the study participants

Of the total study participants, the majority of them (52.7%) were aged between 15 and 19 years old. Regarding educational status, 7.2% did not attend any formal education and nearly half (48.5%) had attended secondary or above school education. 71.8% of participants were rural dwellers. The majority of the study participants (81.5%) were Christian religious followers. Concerning marital status, 60.3% study participants were not married (Table 3).

Pooled prevalence of pregnancy termination in East Africa

The pooled prevalence of pregnancy termination among adolescent and young women (aged 15–24 years old) in East Africa was found to be 4.56%. The highest prevalence was recorded from Uganda (7.6%) followed by Madagascar (7.3%), and Tanzania (5.3%), and the lowest prevalence was reported from Rwanda (2.3%) followed by Ethiopia (2.7%) (Fig. 3).

Feature selection

The Boruta algorithm, respecting the outcome, was utilized to detect the significant feature from the dataset. The variables in the boxplot colored green and arranged by importance in ascending order are greater than the shadowmax and considered relevant by the algorithm. The red-colored variable is smaller than the Shadomax

Variables	Categories	Frequency	Percentage
Participants age	15–19	39,339	52.69
	20–24	35,317	47.31
Age at first sex	< 15	7685	10.29
	≥ 15	66,967	89.71
Educational level	No formal education	5356	7.17
	Primary	33,073	44.30
	Secondary and above	36,227	48.53
Religion	Christian	56,127	81.53
	Muslim	7410	10.76
	Others	5303	7.70
Working status	Not working	41,682	55.84
	Working	32,958	44.16
Wealth index	Poor	26,854	3.97
	Middle	14,158	1.96
	Rich	33,644	45.07
Smoking status	No	68,549	99.60
	Yes	273	0.40
Place of residence	Urban	21,070	28.22
	Rural	53,586	71.78
Distance to HF	Big problem	24,289	34.41
	Not a big problem	44,529	65.59
Marital status	Unmarried	45,025	60.31
	Married	25,825	34.59
	Divorced/widowed/separated	3806	5.10
Mass media exposure	Not exposed	22,282	29.85
	Exposed	52,368	70.15
Parity	No birth	45,349	60.74
	1–4	29,137	39.03
	≥ 5	171	0.23
Children born in the last 5 years	No birth	46,525	62.32
	One	19,872	26.62
	≥ 2	8259	11.06
Community-level educational status	Low	38,292	51.29
	High	36,364	48.71
Community level poverty	High	36,523	48.92
	Low	38,133	51.08
Community-level mass media exposure	Low	36,735	49.21
	High	37,921	50.79

Table 3. Background characteristics of the study participants.

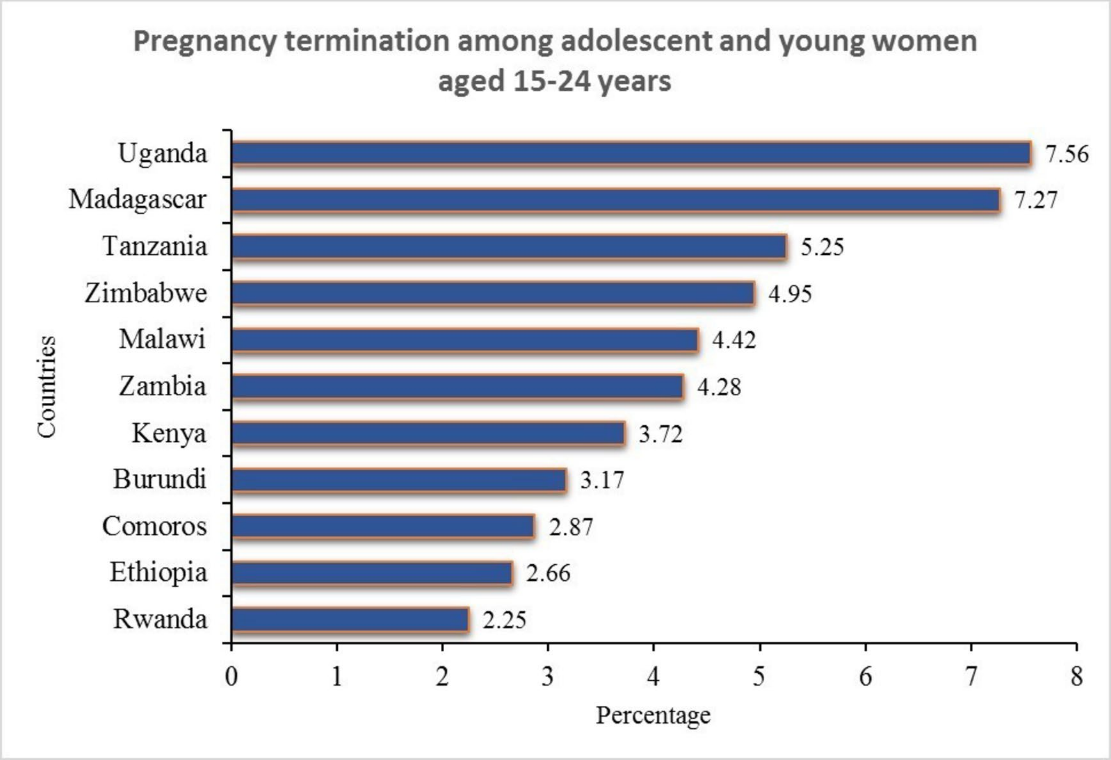

Fig. 3. Prevalence of pregnancy termination among adolescent and young women across East African countries.

and deemed unimportant by the algorithm. Then we excluded smoking status (v06) in our extensive research since the Boruta algorithm considered it to be unimportant. Next, we employed association rule mining to examine data patterns and predicted the pregnancy termination status using the variables we had chosen from the Boruta method. The importance of these factors is illustrated in Fig. 4 below.

Data balancing
Five data balancing techniques namely: under-sampling, oversampling, near miss algorithm, ADASYN, and BorderLine SMOTE were employed and comparison was done accordingly. To determine the optimal data balancing strategy, we performed a thorough examination of the model's performance and considered a number of performance metrics. We trained ML algorithms on the balanced data using multiple balancing procedures and compared their performance. As shown in Table 3, the result indicated that BorderLine SMOTE was the top performer among all the evaluated data balancing techniques. The random forest classifier achieved the highest accuracy (87.0%) and AUC (0.89)notably after data were balanced using BorderLine SMOTE. Hence, for additional research and optimization, we decided to use BorderLine SMOTE. BorderLine SMOTE is a modified version of SMOTE, which emphasizes synthetic sample generation at the boundary between classes[41]. SMOTE uses information about the neighbors who surround each sample in the minority class to construct new artificial instances[42] (Table 4).

Model evaluation and performance comparisons of ML algorithms to predict pregnancy termination
All ML models were fine-tuned by applying the grid search optimization technique, which involves systematically evaluating a set of hyperparameter values to identify the combination that yields the best model performance. Following this comprehensive tuning process, the radial basis function (RBF) kernel was chosen for the Support Vector Machine (SVM) model. This decision was based on the RBF kernel's ability to effectively capture complex patterns in the data, leading to improved classification accuracy compared to other kernel options. Table 5 shows detailed information about the tuned parameter values of each applied ML model.

Nine machine learning algorithms were assessed for their predictive ability to determine how accurately they could predict pregnancy termination based on the performance metrics from the confusion matrix. The models were evaluated based on accuracy, AUC, recall, precision, and F1 score. The accuracy, sensitivity (recall), specificity, and other evaluation metrics values of all included machine learning algorithms are presented in Fig. 5. The evaluation findings showed that the four most effective machine learning models for predicting pregnancy termination were the support vector machine, RF, DT, and CatBoost. RF was the best-performed model with the highest accuracy of 92.9%, this shows that 92.9% of the test set's total data points can be correctly

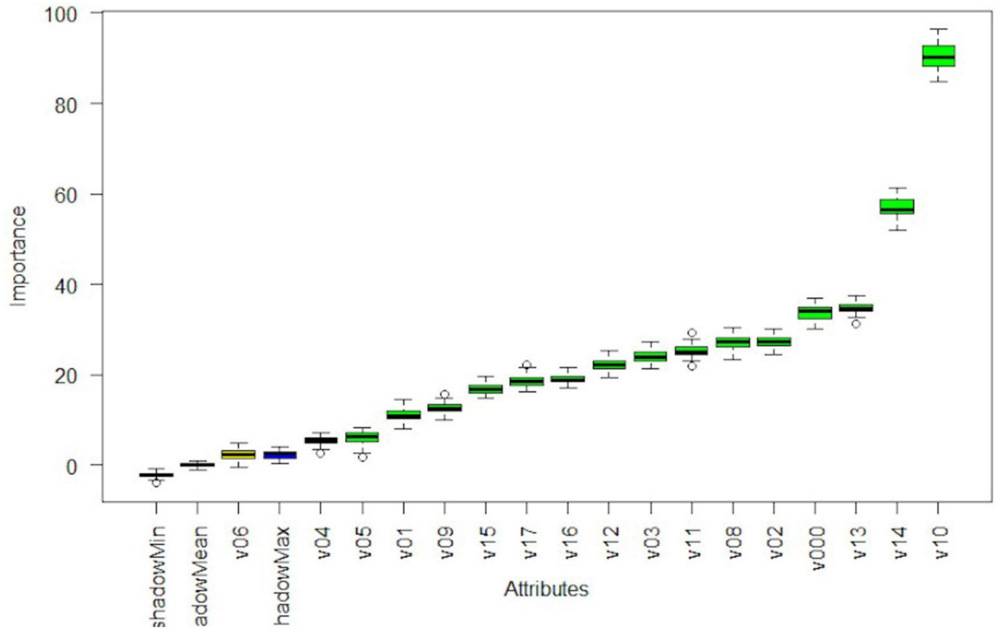

Fig. 4. Feature selection using Boruta algorithm. v01: age, v02: age at first sex, v03: educational level, v04: religion, v05: working status, v06: smoking status, v08: place of residence, v09: distance to HF, v00: country of residence, v10: marital status, v11: wealth index, v12: mass media exposure, v13: number of children ever born, v14: number of births in the last five years, v15: community level educational status, v16: community level poverty, v17: community level mass media exposure.

Models	Performance metrics	Under-sampling	Oversampling	ADASYN	Near miss algorithm	Border-line SMOTE
SVM	AUC	0.78	0.82	0.88	0.82	0.87
	Accuracy	75%	77%	80%	77%	81%
DT	AUC	0.68	0.89	0.85	0.87	0.83
	Accuracy	67%	82%	78%	81%	87%
LR	AUC	0.78	0.78	0.86	0.77	0.81
	Accuracy	73%	73%	76%	72%	76%
RF	AUC	0.76	0.90	0.88	0.88	**0.89**
	Accuracy	70%	83%	81%	82%	**87%**
XGB	AUC	0.77	0.86	0.86	0.85	0.84
	Accuracy	71%	79%	80%	79%	84%
KNN	AUC	0.76	0.84	0.87	0.84	0.86
	Accuracy	77%	77%	78%	77%	78%
AdaBoost	AUC	0.78	0.79	0.85	0.79	0.82
	Accuracy	74%	75%	77%	74%	77%
Catboost	AUC	0.78	0.86	0.90	0.85	0.89
	Accuracy	73%	79%	81%	79%	84%
FNN	AUC	0.78	0.84	0.88	0.84	0.79
	Accuracy	74%	78%	80%	78%	83%

Table 4. Comparison of accuracy and AUC scores across different data balancing techniques. Significant values are in bold.

predicted by our fitted model. DT is the second best-performed model with an accuracy of 87.5% followed by XGB with an accuracy of 86.0% (Fig. 5).

The ROC plot for all nine machine learning models is shown in Fig. 6. As shown in the figure, the constructed prediction models had different predictive performances. RF performed the highest with an AUC value of 0.93 followed by DT, XGB, and CatBoost having the same AUC values of 0.92.

ML models	Accuracy	AUC	Sensitivity	Specificity	F1-score
SVM	0.816	0.88	0.907	0.778	0.851
DT	0.875	0.92	0.931	0.854	0.882
LR	0.722	0.81	0.818	0.662	0.778
RF	0.929	0.93	0.971	0.874	0.919
XGB	0.860	0.92	0.921	0.787	0.843
KNN	0.718	0.91	0.939	0.716	0.847
AdaBoost	0.764	0.82	0.877	0.677	0.797
Catboost	0.859	0.92	0.93	0.791	0.875
FNN	0.818	0.90	0.903	0.741	0.839

Table 5. The tuned parameter values of each applied ML model.

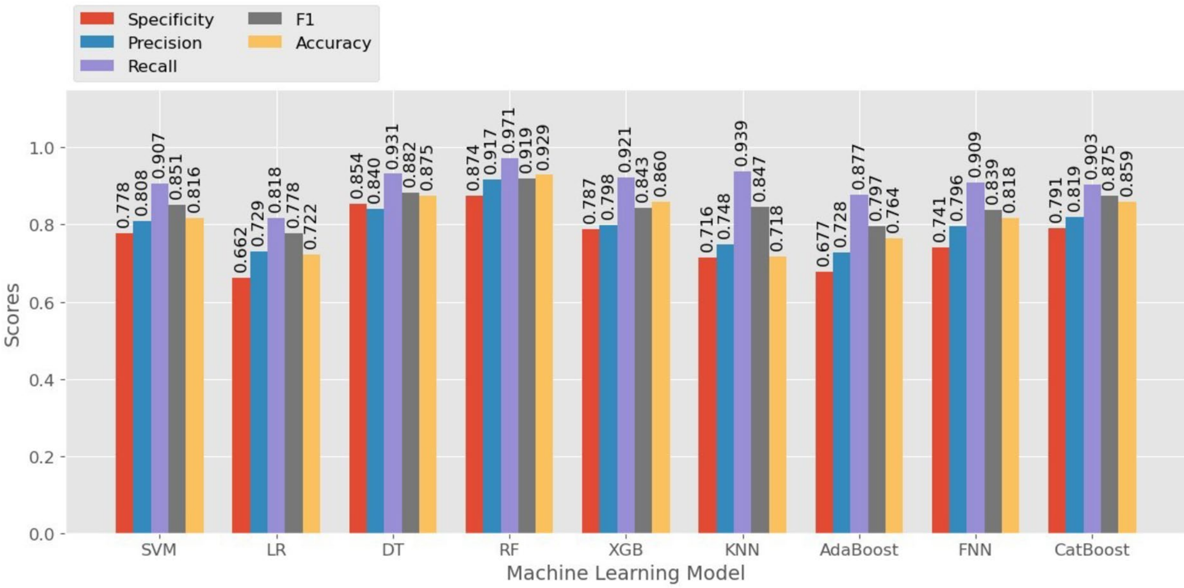

Fig. 5. Model performance of each ML model after optimized Hyperparameter tuning.

Model interpretability

Random forest based feature importance

Important factors that influence pregnancy termination were measured. The best performance model, random forest, was considered in order to choose important predictors. Accordingly, the most significant variables of pregnancy termination were country of residence, marital status, wealth index, educational attainment, distance to HF, religion, and mass media exposure. Important attributes of pregnancy termination are illustrated in Fig. 7.

SHAP feature impact on model prediction

The SHAP (SHapley Additive exPlanations) analysis presented in Fig. 8 reveals the feature importance scores for predicting pregnancy termination. The features are listed on the vertical axis in descending order of their impact on the model's predictions. The most important feature is at the top, and the least important is at the bottom. The horizontal axis represents the SHAP values, which indicate the impact of each feature on the model's output. Positive values suggest that the feature increases the likelihood of the predicted outcome (pregnancy termination), while negative values indicate a decrease. Features with a significant number of points on the left (negative SHAP values) suggest that they are associated with lower likelihoods of the predicted outcome, while those on the right (positive SHAP values) indicate higher likelihoods. The findings highlight marital status, age, parity, country of residence, age at first sexual activity, exposure to mass media, and educational status as the most significant influencing factors for predicting pregnancy termination (Fig. 8).

Figure 9 presents a SHAP force plot that illustrates the impact of each feature on a model's prediction using SHAP values. This visualization clarifies the positive and negative contributions of different features to the model's output. As shown in the figure, features that push the model towards predicting a higher pregnancy termination are represented on the left in red, whereas those features that push the model towards a lower pregnancy termination are shown on the right in blue. Features with large SHAP values have larger arrows.

Hence, The SHAP values for age at first sex show that lower values (age at first sex below 15 years old) are associated with a positive impact on the output, indicating that younger age at first sexual activity increases

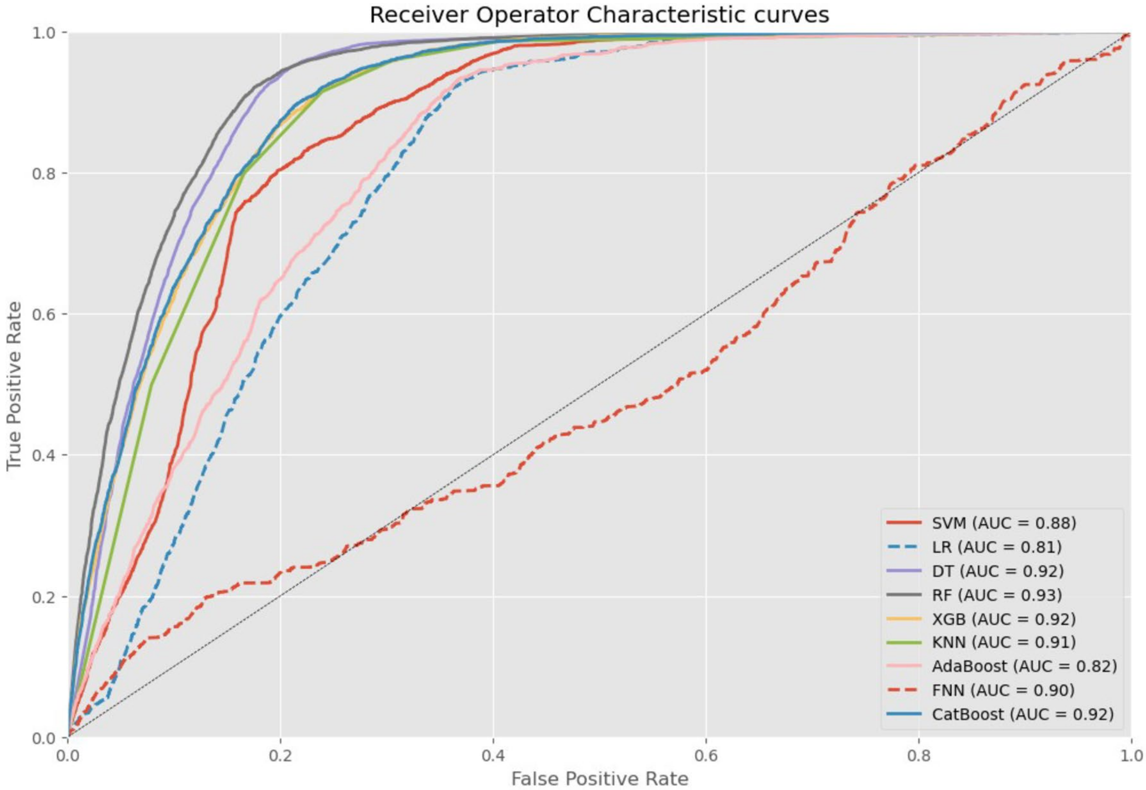

Fig. 6. ROC curve analysis of each ML model after optimized Hyperparameter tuning.

the likelihood of the predicted outcome (pregnancy termination). The SHAP values of marital status indicate that being divorced/separated has a positive impact on the predicted outcome, suggesting that being divorced/separated is associated with an increased likelihood of pregnancy termination. Lower values of age (15–19 years) have a positive impact on the predicted outcome, indicating that younger age is associated with an increased likelihood of pregnancy termination. However, lower parity values tend to have a negative impact on the predicted outcome, suggesting that having lower parity decreases the likelihood of pregnancy termination (Fig. 9).

Discussion

Machine learning is vital in solving many medical and public health problems and offers a wide range of applications[43]. Hence, this study investigates various machine learning algorithms to predict pregnancy termination and find important features linked with it. We utilized the most recent DHS datasets, which included from eleven East African countries. The pooled prevalence of pregnancy termination in East Africa among adolescent and young women aged 15 to 24 was 4.56%.

To the best of our knowledge, this effort is the first attempt to use various machine learning techniques to predict pregnancy termination using data from East African countries. The study's main objective was to assess machine learning techniques and identify the best model for predicting pregnancy termination. Accordingly, nine machine learning models were considered and evaluation was done based on different aspects of the evaluation metrics. Five-fold cross-validation was used to train each model using the training data set, and the test data set was used to evaluate performance.

This study demonstrated that three models—random forest, decision tree, and XGB performed the highest prediction accuracies and AUC statistics in terms of predicting pregnancy termination. However, the RF predictive model performs better than other predictive models when comparing across various evaluation metrics. In terms of the predictive analysis, the prediction accuracy of the random forest was 92.9%%, the AUC was 0.93, the specificity was 0.87, and the precision was 0.91.

Previous studies used a variety of machine learning techniques to predict pregnancy outcomes. Upon analyzing the algorithms' performance, distinct algorithms showed different results in terms of certain model performance indicators. For e.g. Amitai T., et al.[44] achieved an AUC of 0.69 for predicting miscarriage outcomes. The results of this study are interestingly consistent with our research since they showed that the random forest model outperformed the other models. Consistent with our findings, comparable studies have indicated that RF is the best-performing model for predicting low birth weight[45], early pregnancy loss[46], and preterm birth in women with cervical cerclage[47].

While Wu Y., et al.[48] got an AUC of 0.9209 with an accuracy of 84.69% using the XGBoost model for predicting the risk of miscarriage among immune abnormal pregnancies. Wu Y., et al. reportedly outperformed XGBoost in

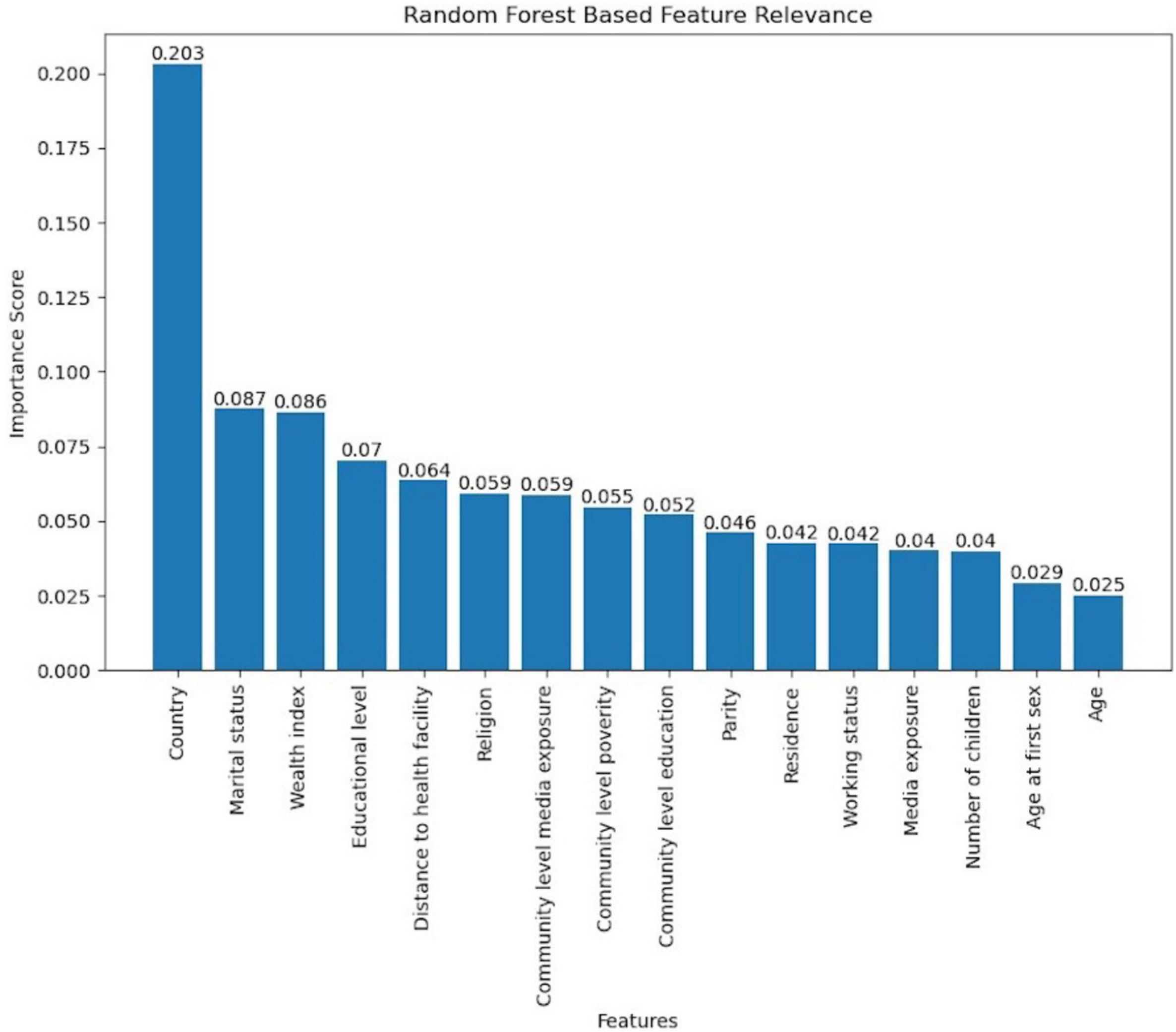

Fig. 7. Random forest-based feature importance.

terms of prediction accuracy. However, bb used small sample size ($n = 565$) to develop and evaluate ML models. Nevertheless, our analysis revealed that RF exhibited superior performance compared to other models. Our study utilized multi-country large datasets and encompassed a wider array of algorithms.

The other objective of this study was to identify important characteristics that might be used to predict pregnancy termination. According to the insights derived from the SHAP analysis; marital status, age, parity, country of residence, age at first sex, mass media exposure, and educational status are the most influential factors for predicting pregnancy termination among adolescent and young women.

This study showed that marital status is an important feature for predicting pregnancy termination among adolescent and young women aged 15 to 24 years. This finding is supported by various studies done in Nigeria[49,50], Ethiopia[51], Sierra Leone[52], Burkina Faso[53], Ghana[54], and Brazil[55], which revealed that single women (widowed/separated/divorced and never married) were more likely to undergo a pregnancy termination than married women. If adolescents and young women are not in marital relationships, the chance of being exposed to unwanted pregnancy will be high. Additionally, the increased risk of pregnancy termination among divorced or widowed women may be due to their desire to prevent additional pregnancies. Moreover, if a widowed woman becomes pregnant, she may encounter criticism and potential social exclusion, most African nations stigmatize unmarried mothers who have children could be a possible reason.

Similar to marital status, age appears to be a crucial factor for predicting pregnancy termination, with varying impacts based on the specific age. In line with this, previous studies have shown that age can be a significant factor in pregnancy termination[56]. Younger adolescents may be more likely to face barriers in accessing reproductive health services. Additionally, younger age is often associated with higher rates of unintended pregnancies, which can also lead to higher rates of pregnancy termination[57].

Based on the findings of this study, the other most important variable linked to pregnancy termination was the number of children ever born, which is supported by various studies[56,58–61]. Women who have decided on their desired family size may choose to end an unintended pregnancy to control the size of their family[60,62].

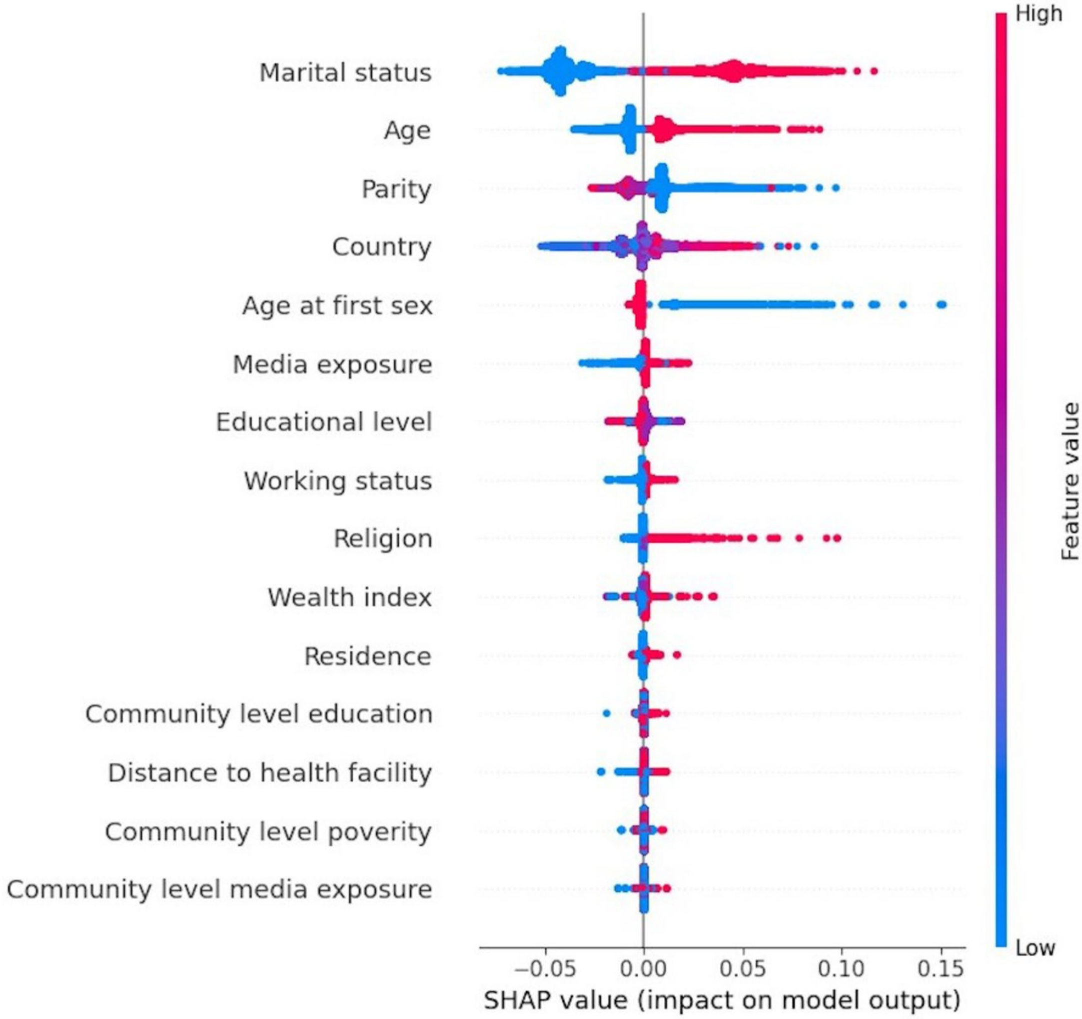

Fig. 8. SHAP feature impact on model prediction.

Fig. 9. SHAP force plot.

The SHAP values suggest that country of residence is one of the most important variables in predicting pregnancy termination among adolescent and young women in East African countries. Figure 3 also clearly demonstrates significant variations in pregnancy termination across different countries in our descriptive analysis. Previous studies also show that country variation does have a significant impact on pregnancy termination rates. For example, estimates from the WHO and the Guttmacher Institute shows significant differences in the rates of unwanted pregnancies and abortions between different countries[63]. In Europe even, the rates of abortion and unwanted pregnancies varied, while in sub-Saharan Africa, they are widely varied[63]. Another study also reported regional differences in the prevalence of pregnancy termination in sub-Saharan Africa[64]. Factors such as access to healthcare, legal restrictions, availability of contraception, and cultural attitudes may contribute to these differences. Moreover, the reasons for the varied rates of pregnancy termination could be due to the influence of complex interplay factors such as access to education[29], social support systems[65], religiosity[60,61,66,67], and economic factors[12].The findings of this study showed that educational attainment is also the most important predictor of pregnancy termination. Many studies have also highlighted the crucial impact of this variable[52,56,62,68]. Women with higher levels of education are more aware of the legislation and abortion service providers[69]. Additionally, educated women are more likely to utilize contraception to avoid undesired pregnancies, which are more common and can result in pregnancy termination[57].

This study shows that mass media exposure is an important attribute for predicting pregnancy termination. Previous studies have demonstrated a strong link between the abortion self-efficacy of women in their reproductive age group and their media exposure[70–72]. A woman familiar with the mass media may know how and where to access pregnancy termination services. Moreover, such individuals may also possess knowledge of existing abortion laws and may encounter less social stigma[73].

Limitations
It is important to take into account the limitations of this study. Firstly, the dataset used in this study came from secondary data sources, which have limitations in the inclusion of certain possible clinical features for predicting pregnancy termination. Therefore, it is highly recommended to conduct additional external validations utilizing data from multicenter settings. Additionally, it is important to recognize that the variables in this study are derived from self-reported data, which can be subject to recall bias and social desirability bias, potentially affecting the accuracy of the responses. Hence, in future studies, it would be helpful to employ measures that are more objective or triangulate self-reported data with other data sources to enhance the validity and reliability of results.

Conclusions and recommendations
In this study, nine state-of-the-art machine learning models, namely SVM, LR, RF, DT, XGB, KNN, AdaBoost, FNN, and CatBoost were constructed and evaluated. The RF model outperformed all the other models based on a number of evaluation metrics. Marital status, age, parity, country of residence, age at first sexual activity, exposure to mass media, and educational attainment were found the most important attributes for predicting pregnancy termination.

This study has significant methodological and practice implications. In terms of methodology, we applied cutting-edge machine learning models to nationally representative data from eleven datasets in East Africa, and this allowed us to identify important predictive features. In practice, the study identified a set of important factors that can be used to accurately predict the risk of pregnancy termination for young and adolescent women, enabling more targeted interventions.

Therefore, the results of the study would be useful to maternal and child health programmers, decision-makers, and other stakeholders when developing interventions for mothers and children to reduce morbidity and mortality. Taking into account these important variables would help to tackle the problem. Additionally, it is important to have involvement from multiple sectors, including religious and community leaders, to raise awareness about pregnancy termination and its implications. Furthermore, this study would serve as a baseline for future studies on related topics, encompassing important factors that could predict pregnancy termination. The recommended model has the best ROC value, accuracy, sensitivity, precision, Fl score, and specificity for predicting pregnancy termination. Furthermore, machine learning models play a crucial role in predicting health studies that lead to improved and more suitable policy decisions.

Data availability
The dataset used and/or analyzed during the current study is available from the corresponding author upon reasonable request.

Received: 3 March 2024; Accepted: 25 November 2024
Published online: 03 December 2024

References
1. Harvard Health Publishing. Abortion (Termination Of Pregnancy): what is it? https://www.health.harvard.edu/medical-tests-and-procedures/abortion-termination-of-pregnancy-a-to-z (2019).
2. Clark Alves, S. M. October, Jenkins, Amanda Rapp. Early Pregnancy Loss (Spontaneous Abortion), 12 (StatPearls Publishing, 2023).
3. Diedrich, J. & Steinauer, J. Complications of surgical abortion. *Clin. Obstet. Gynecol.* **52**, 205–212 (2009). https://doi.org/10.1097/GRF.0b013e3181a2b756.
4. Grimes, D. A. et al. Unsafe abortion: the preventable pandemic. *Lancet* **368**, 1908–1919. https://doi.org/10.1016/S0140-6736(06)69481-6 (2006).
5. United Nations Population Fund. Seeing the unseen: The case for action in the neglected crisis of unintended pregnancy. https://www.unfpa.org/swp2022 (2022).
6. Say, L. et al. Global causes of maternal death: a WHO systematic analysis. *Lancet Global Health* **2**, e323–e333. https://doi.org/10.1016/S2214-109X(14)70227-X (2014).
7. Ganatra, B. et al. Global, regional, and subregional classification of abortions by safety, 2010–14: estimates from a Bayesian hierarchical model. *Lancet* **390**, 2372–2381. https://doi.org/10.1016/S0140-6736(17)31794-4 (2017).
8. Okonofua, F. Abortion and maternal mortality in the developing world. *J. Obstet. Gynecol. Can.* **28**, 974–979. https://doi.org/10.1016/S1701-2163(16)32307-6 (2006).
9. World Health Organization. Strategies towards ending preventable maternal mortality (EPMM, 2015).
10. Allotey, P., Ravindran, T. K. S. & Sathivelu, V. Trends in abortion policies in low- and middle-income countries. *Annu. Rev. Public Health* **42**, 505–518. https://doi.org/10.1146/annurev-publhealth-082619-102442 (2021).
11. Frederico, M. et al. Factors influencing abortion decision-making processes among young women. *Int. J. Environ. Res. Public Health* **15**. https://doi.org/10.3390/ijerph15020329
12. Gbagbo, F. Y., Amo-Adjei, J. & Laar, A. Decision-making for induced abortion in the Accra metropolis, Ghana. *Afr. J. Reprod. Health* **19**, 34–42 (2015).

13. Sharifi, A., Janatolmakan, M. & Khatony, A. The prevalence and the reasons of issuing permission for therapeutic abortion in department of forensic medicine, Kermanshah, Iran, during 2005 to 2010. *BMC Res. Notes* **12**, 574. https://doi.org/10.1186/s13104-019-4622-4 (2019).
14. Secinaro, S. et al. The role of artificial intelligence in healthcare: a structured literature review. *BMC Med. Inf. Decis. Mak.* **21**, 125. https://doi.org/10.1186/s12911-021-01488-9 (2021).
15. Tagliaferri, S. D. et al. Artificial intelligence to improve back pain outcomes and lessons learnt from clinical classification approaches: three systematic reviews. *Npj Digit. Med.* **3**, 93. https://doi.org/10.1038/s41746-020-0303-x (2020).
16. Cho, B-J. et al. Classification of cervical neoplasms on colposcopic photography using deep learning. *Sci. Rep.* **10**, 13652. https://doi.org/10.1038/s41598-020-70490-4 (2020).
17. Tran, B. X. et al. Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. *J. Clin. Med.* **8**. https://doi.org/10.3390/jcm8030360
18. Panch, T., Szolovits, P. & Atun, R. Artificial intelligence, machine learning and health systems. *J. Glob. Health* **8**, 020303. https://doi.org/10.7189/jogh.08.020303 (2018).
19. Yang, X. et al. Concepts of artificial intelligence for computer-assisted drug discovery. *Chem. Rev.* **119**, 10520–10594. https://doi.org/10.1021/acs.chemrev.8b00728 (2019).
20. Javaid, M. et al. Significance of machine learning in healthcare: features, pillars and applications. *Int. J. Intell. Netw.* **3**, 58–73. https://doi.org/10.1016/j.ijin.2022.05.002 (2022).
21. Jay, N. et al. A data mining approach for grouping and analyzing trajectories of care using claim data: the example of breast cancer. *BMC Med. Inf. Decis. Mak.* **13**, 130. https://doi.org/10.1186/1472-6947-13-130 (2013).
22. Aljumah, A. A., Ahamad, M. G. & Siddiqui, M. K. Application of data mining: diabetes health care in young and old patients. *J. King Saud Univ. Comput. Inform. Sci.* **25**, 127–136. https://doi.org/10.1016/j.jksuci.2012.10.003 (2013).
23. Delen, D., Walker, G. & Kadam, A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artif. Intell. Med.* **34**, 113–127. https://doi.org/10.1016/j.artmed.2004.07.002 (2005).
24. Chen, H-Y. et al. Exploring the risk factors of preterm birth using data mining. *Expert Syst. Appl.* **38**, 5384–5387. https://doi.org/10.1016/j.eswa.2010.10.017 (2011).
25. Heikinheimo, O., Gissler, M. & Suhonen, S. Age, parity, history of abortion and contraceptive choices affect the risk of repeat abortion. *Contraception* **78**, 149–154. https://doi.org/10.1016/j.contraception.2008.03.013 (2008).
26. Saseendran, P. & Stones, R. W. Maternal and social factors associated with abortion in India: a population-based study. *Int. Fam. Plan. Perspect.* **32**, 120–125 (2006).
27. Koch, E. et al. Women's education level, maternal health facilities, abortion legislation and maternal deaths: a natural experiment in Chile from 1957 to 2007. *PLoS One* **7**, e36613. https://doi.org/10.1371/journal.pone.0036613 (2012).
28. Andersen, K. L. et al. Marital status and abortion among young women in Rupandehi, Nepal. *BMC Women's Health* **15**, 17. https://doi.org/10.1186/s12905-015-0175-4 (2015).
29. Ba, D. M. et al. Factors associated with pregnancy termination in women of childbearing age in 36 low-and middle-income countries. *PLOS Glob. Public Health* **3**, e0001509. https://doi.org/10.1371/journal.pgph.0001509 (2023).
30. Dai, R. et al. Effect of maternal age on spontaneous abortion during the first trimester in Northeast China. *J. Maternal-Fetal Neonatal Med.* **31**, 1824–1829. https://doi.org/10.1080/14767058.2017.1330330 (2018).
31. Chen, R-C. et al. Selecting critical features for data classification based on machine learning methods. *J. Big Data* **7**, 52. https://doi.org/10.1186/s40537-020-00327-4 (2020).
32. Hornik, K. Classification and regression by randomForest 18. Some strategies for dealing with genomic data 22. Changes to the R-Tcl/Tk package. https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf
33. Stańczyk, U. & Jain, L. C. Feature selection for data and pattern recognition: an introduction. In: *Feature Selection for Data and Pattern Recognition* (eds Stańczyk, U. & Jain, L. C.), 1–7 (Springer Berlin Heidelberg, 2015).
34. Arafat, M. Y. et al. Machine learning for mining imbalanced data (2018).
35. Chawla, N. V. Data Mining for imbalanced datasets: an overview. In: *Data Mining and Knowledge Discovery Handbook* (eds Maimon, O. & Rokach, L.), 853–867 (Springer US, 2005).
36. Batista, G. E. A. P. A., Prati, R. C. & Monard, M. C. A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor. Newsl.* **6**, 20–29. https://doi.org/10.1145/1007730.1007735 (2004).
37. Rosellini, A. J. et al. Developing algorithms to predict adult onset internalizing disorders: an ensemble learning approach. *J. Psychiatr. Res.* **121**, 189–196. https://doi.org/10.1016/j.jpsychires.2019.12.006 (2020).
38. Florkowski, C. M. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin. Biochem. Rev.* **29** (Suppl 1), S83–87 (2008).
39. Mitra, S., Pal, S. K. & Mitra, P. Data mining in soft computing framework: a survey. *IEEE Trans. Neural Netw.* **13**, 3–14 (2002).
40. Hosmer, D. W. Jr, Lemeshow, S. & Sturdivant, R. X. *Applied Logistic Regression* (Wiley, 2013).
41. Han, H., Wang, W-Y., Mao, B-H. & Borderline -SMOTE: A new over-sampling method in imbalanced data sets learning. In: *Advances in Intelligent Computing* (eds Huang D-S, Zhang X-P and Huang G-B), 878–887 (Springer Berlin Heidelberg, 2005).
42. Nnamoko, N. & Korkontzelos, I. Efficient treatment of outliers and class imbalance for diabetes prediction. *Artif. Intell. Med.* **104**, 101815. https://doi.org/10.1016/j.artmed.2020.101815 (2020).
43. Maity, N. G. & Das, S. March. Machine learning for improved diagnosis and prognosis in healthcare. In: *2017 IEEE Aerospace Conference*, 1–9 (2017).
44. Tamar Amitai, Y. K. T. et al. Embryo classification beyond pregnancy: early prediction of first trimester miscarriage using machine learning. *J. Assist. Reprod. Genet.* **40**, 309–322. https://doi.org/10.1007/s10815-022-02619-5 (2023).
45. Ghaderighahfarokhi, S., Sadeghifar, J. & Mozafari, M. A model to predict low birth weight infants and affecting factors using data mining techniques. *J. Basic Res. Med. Sci.* **5**, 1–8. https://doi.org/10.29252/jbrms.5.3.1 (2018). Research.
46. Lijue, L. et al. Machine learning algorithms to predict early pregnancy loss after in vitro fertilization-embryo transfer with fetal heart rate as a strong predictor. *Comput. Methods Progr. Biomed.* **196**, 105624. https://doi.org/10.1016/j.cmpb.2020.105624 (2020).
47. Hasan, R. et al. Intelligent system based on data mining techniques for prediction of preterm birth for women with cervical cerclage. *Comput. Biol. Chem.* **85**, 107233. https://doi.org/10.1016/j.compbiolchem.2020.107233 (2020).
48. Wu, Y. et al. Risk prediction model based on machine learning for predicting miscarriage among pregnant patients with immune abnormalities. *Front. Pharmacol.* **15**. https://doi.org/10.3389/fphar.2024.1366529 (2024). Original Research.
49. Ibrahim, I. A. & Onwudiegwu, U. Sociodemographic determinants of complicated unsafe abortions in a semi-urban Nigerian town: a four-year review. *West Indian Med. J.* **61** (2012).
50. Mitsunaga, T. M., Larsen, U. M. & Okonofua, F. E. Risk factors for complications of induced abortions in Nigeria. *J. Women's Health* **14**, 515–528. https://doi.org/10.1089/jwh.2005.14.515 (2005).
51. Gilano, G. & Hailegebreal, S. Determinants of abortion among youth 15–24 in Ethiopia: a multilevel analysis based on EDHS 2016. *PLoS One* **16**, e0248228. https://doi.org/10.1371/journal.pone.0248228 (2021).
52. Sesay, F. R. et al. Determinants of induced abortion among women of reproductive age: evidence from the 2013 and 2019 Sierra Leone Demographic and Health Survey. *BMC Women's Health* **23**, 44. https://doi.org/10.1186/s12905-023-02175-9 (2023).
53. Ilboudo, P. G. C., Somda, S. M. A. & Sundby, J. Key determinants of induced abortion in women seeking postabortion care in hospital facilities in Ouagadougou, Burkina Faso. *Int. J. Women's Health*. **6**, 565–572. https://doi.org/10.2147/IJWH.S60709 (2014).
54. Klutsey, E. E. & Ankomah, A. Factors associated with induced abortion at selected hospitals in the Volta Region, Ghana. *Int. J. Women's Health* **6**, 809–816. https://doi.org/10.2147/IJWH.S62018 (2014).

55. Misago, C. et al. Determinants of abortion among women admitted to hospitals in Fortaleza, North Eastern Brazil. *Int. J. Epidemiol.* **27**, 833–839 (1998).
56. Megersa, B. S. et al. Factors associated with induced abortion among women of reproductive age attending selected health facilities in Addis Ababa, Ethiopia: a case control study. *BMC Women's Health* **20**, 188. https://doi.org/10.1186/s12905-020-01023-4 (2020).
57. Ahinkorah, B. O. Socio-demographic determinants of pregnancy termination among adolescent girls and young women in selected high fertility countries in sub-saharan Africa. *BMC Pregnancy Childbirth* **21**, 598. https://doi.org/10.1186/s12884-021-04064-1 (2021).
58. Tesfaye, G., Hambisa, M. T. & Semahegn, A. Induced abortion and associated factors in health facilities of Guraghe Zone, Southern Ethiopia. *J. Pregnancy* **2014**, 295732. https://doi.org/10.1155/2014/295732 (2014).
59. Tsakiridu, D. O. et al. Factors associated with induced abortion in women prostitutes in Asturias (Spain). *PLoS One* **3**, e2358. https://doi.org/10.1371/journal.pone.0002358 (2008).
60. Ranji, A. Induced abortion in Iran: prevalence, reasons, and consequences. *J. Midwifery Women's Health* **57**, 482–488. https://doi.org/10.1111/j.1542-2011.2012.00159.x (2012).
61. Erfani, A. Induced abortion in Tehran, Iran: estimated rates and correlates. *Int. Perspect. Sex. Reproductive Health* **37**, 134–142 (2011).
62. Sihvo, S. et al. Women's life cycle and abortion decision in unintended pregnancies. *J. Epidemiol. Commun. Health* **57**, 601. https://doi.org/10.1136/jech.57.8.601 (2003).
63. WHO. First ever country level estimates of unintended pregnancy and abortion. https://www.who.int/news/item/24-03-2022-first-ever-country-level-estimates-of-unintended-pregnancy-and-abortion (2022).
64. Rahel Nega Kassa, E. W. K., Uduak Andrew-Bassey, S. A., Abdiwali, B. & Nahayo Ndayishim ye Samuel, Joshua Odunayo Akinyemi. Factors associated with pregnancy termination in six sub-saharan African countries. *PLoS Glob. Public Health* **4** https://doi.org/10.1371/journal.pgph.0002280 (2024).
65. Sorhaindo, A. M. et al. Qualitative evidence on abortion stigma from Mexico City and five states in Mexico. *Women Health* **54**, 622–640. https://doi.org/10.1080/03630242.2014.919983 (2014).
66. Yogi, A., Neupane, S. & K.C P and Prevalence and factors associated with abortion and unsafe abortion in Nepal: a nationwide cross-sectional study. *BMC Pregnancy Childbirth* **18**, 376. https://doi.org/10.1186/s12884-018-2011-y (2018).
67. Santos APVd, Coelho, E. A. C. et al. Factors Associated with abortion in women of reproductive age. *Revista Brasileira De Ginecol. e Obstetrícia* **38** (2016).
68. Xu, J. S., Huang, Y. M. & Cheng, L. N. [Factors in relation to repeated abortions among unmarried young people in Shanghai]. *Zhonghua Liu Xing Bing Xue Za Zhi*. **28**, 742–745 (2007).
69. Yaya, S. et al. Women empowerment as an enabling factor of contraceptive use in sub-saharan Africa: a multilevel analysis of cross-sectional surveys of 32 countries. *Reprod. Health* **15**, 214. https://doi.org/10.1186/s12978-018-0658-5 (2018).
70. Tesema, G. A., Mekonnen, T. H. & Teshale, A. B. Spatial distribution and determinants of abortion among reproductive age women in Ethiopia, evidence from Ethiopian Demographic and Health Survey 2016 data: spatial and mixed-effect analysis. *PLoS One* **15**, e0235382. https://doi.org/10.1371/journal.pone.0235382 (2020).
71. Dickson, K. S., Adde, K. S. & Ahinkorah, B. O. Socio – economic determinants of abortion among women in Mozambique and Ghana: evidence from demographic and health survey. *Archives of Public Health* **76**, 37. https://doi.org/10.1186/s13690-018-0286-0.72 (2018).
72. Ahinkorah, B.O. et al. Mass media exposure and self-efficacy in abortion decision-making among adolescent girls and young women in Ghana: Analysis of the 2017 Maternal Health Survey. *PL0S One* **15**, e0239894. https://doi.org/10.1371/journal.pone.0239894 (2020).
73. Cook, R. J. & Dickens, B. M. Reducing stigma in reproductive health. *Int. J. Gynecol. Obstet.* **125**, 89–92. https://doi.org/10.1016/j.ijgo.2014.01.002 (2014).

Acknowledgements
The authors would like to thank Measure Demographic and Health Survey for granting permission to use the data from East African countries for this research.

Author contributions
GAT and HSN - Conceptualization, Data curation, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editingADD, BTS, MDK, and TMT - Conceptualization, Data curation, Investigation, Methodology, Resources, Writing – review & editingSH & BTG - Conceptualization, Data curation, Methodology, Project administration, Software, Supervision, Validation, Writing – review & editingAll authors contributed significantly. Additionally, after reading the draft and making any necessary critical revisions for significant intellectual content, all authors approved the work in its final form.

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval and consent to participate
Since this study was done using a secondary data source, participants' consent is not applicable. However, a permission letter and approval were obtained from DHS which allows us to use the data.

Additional information
Correspondence and requests for materials should be addressed to G.A.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.