Original Paper

Identifying Predictors of Cervical Cancer Screening Uptake in Sub-Saharan Africa Using Machine Learning: Cross-Sectional Study

Nebebe Demis Baykemagn¹, MPH; Mekuriaw Nibret Aweke², MPH; Amare Mesfin², MPH; Lemlem Daniel Baffa², MPH; Muluken Chanie Agimas³, MPH; Habtamu Wagnew Abuhay³, MPH; Dagnew Getnet Adugna⁴, MSc; Tewodros Getaneh Alemu⁵, MSc; Alemu Teshale Bicha⁶, MD, MPH; Gebrie Getu Alemu³, MPH

Corresponding Author:

Nebebe Demis Baykemagn, MPH Department of Health Informatics, Institute of Public Health College of Medicine and Health Sciences, University of Gondar Gondar 196 Ethiopia

Phone: 251 918310345 Email: nebebe2@gmail.com

Abstract

Background: Cervical cancer has been ranked as the fourth most common cancer affecting women, contributing to approximately 660,000 new diagnoses and 350,000 fatalities worldwide. Effective early screening has been shown to reduce cervical cancer incidence by up to 80% and prevent more than 40% of new cases.

Objective: This study aims to assess a machine learning–based prediction model and identify the key predictors influencing cervical cancer screening uptake among women aged 30-49 years in sub-Saharan Africa.

Methods: For this study, a weighted dataset of 33,952 individuals from the 2022 Demographic and Health Survey in Ghana, Kenya, Mozambique, and Tanzania was used. STATA version 17 (StataCorp) and Python 3.10 (Python Software Foundation) were used for data preprocessing and analysis. MinMax and standard scaler were applied for feature scaling, and recursive feature elimination was used for feature selection. An 80:20 ratio was applied for data splitting. Tomek links with random oversampling were used for handling class imbalance. A total of 7 models were selected and trained using both balanced and unbalanced datasets. Model evaluation was performed using area under the receiver operating characteristic curve, accuracy, and a confusion matrix.

Results: The proportion of cervical cancer screening in sub-Saharan Africa was 13%, which is lower than reported in previous studies. Random forest was the best-performing model, achieving an accuracy of 78%, an area under the curve of 86%, an F_1 -score of 79%, a recall of 81%, and a precision of 77%. The waterfall plot's Shapley Additive Explanations analysis showed that wealth status, awareness of sexually transmitted infections, HIV testing exposure, age at first sexual intercourse, educational level, residency, smartphone ownership, having a single sexual partner, and previous health status were predictors of cervical cancer screening.

Conclusions: Improving education and awareness, expanding access to screening (especially in rural areas), leveraging both digital health and community-based outreach, integrating screening with other health services, and addressing socioeconomic barriers are recommended strategies to increase cervical cancer screening rates in sub-Saharan Africa.

¹Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia

²Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

³Department of Epidemiology and Biostatistics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

⁴Department of Human Anatomy, School of Medicine, College of Medicine and Health Science, University of Gondar, Gondar, Ethiopia

⁵Department of Pediatrics and Child Health Nursing, School of Nursing, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

⁶Department of Obstetrics and Gynecology, School of Medicine, College of Medicine and Health Sciences, University of Gondar, Ethiopia

JMIR Public Health Surveill 2025;11:e71677; doi: 10.2196/71677

Keywords: cervical cancer; digital health; machine learning; woman; artificial intelligence

Introduction

Cervical cancer has been ranked as the fourth most common cancer affecting women, contributing to approximately 660,000 new diagnoses and 350,000 fatalities globally [1]. About 75%-85% of these cases occur in low-income countries, with 23% specifically in Africa (as reported by the World Health Organization [WHO]). The incidence and mortality rates of cervical cancer are 2-3 times higher in low-income countries compared to high-income countries [2].

Cervical cancer ranks as the second leading cause of cancer-related death globally [3]. In 2025, studies from Africa showed that cervical cancer screening prevalence differs significantly between high-income countries (84%) and lowand middle-income countries (15%) [4]. Also, in Africa, more than 53% of cervical cancer cases are identified at an advanced stage (stage III-IV) [5].

Cervical cancer is "a malignant tumor that forms when the cells in the tissue surrounding the cervix grow and multiply uncontrollably by passing the normal process of cell division" [6]. Effective early screening has been shown to reduce cervical cancer incidence by up to 80% and prevent more than 40% of new cases [2,7].

With an average cure rate of 80%, radiation therapy effectively treats cervical cancer in its early stages [8]. However, the practice of early screening, before the disease reaches an advanced stage, continues to be a major problem in many parts of Africa [5].

Despite the implementation of screening initiatives and the introduction of the human papillomavirus (HPV) vaccine, cervical cancer remains a significant public health concern in Africa [9].

Early detection and timely treatment of cervical cancer are crucial [10]. The HPV has more than 100 serotypes, but the HPV vaccine targets only 9 of them [11]. Also, early screening is a key approach to achieving 1 of the 3 WHO 90-70-90 pillars: "90% of girls fully vaccinated against HPV, 70% of women screened by age 35 and 45 with a high-performance test, and 90% of women with precancerous lesions receiving treatment" [12,12].

Therefore, early screening is essential to enhance the probability of early diagnosis and treatment. Early screening is guaranteed to identify precancerous lesions and cancer before symptoms appear, which is the goal of screening, and enhances the chances of successful treatment and reduces maternal mortality [12-14].

However, studies show that screening in Africa is conducted at less than 70% of the WHO target [15]. This is due to a perceived lack of expertise, limited knowledge about cervical cancer, inadequate screening tools and supplies, and insufficient funding [5,16,17]. In the recent study, the prevalence of precancerous cervical lesions and cancer is relatively high, at 18% and 14%, respectively [18].

According to previous studies conducted in Africa, age, education level, employment status, wealth index, place of residence, and access to health care facilities were common predictors of cervical cancer screening among women of reproductive age [19-22].

Although studies have been conducted in African countries, they were limited to small sample sizes and single countries, lacking representativeness. This study, however, is representative, with a large dataset covering multiple countries, providing reliable evidence for policymakers and researchers to inform interventions.

The health care system produces vast amounts of data that must be analyzed to support effective decision-making. As the volume and complexity of this data grow, relying on traditional analytical methods without the aid of machine learning is becoming increasingly impractical.

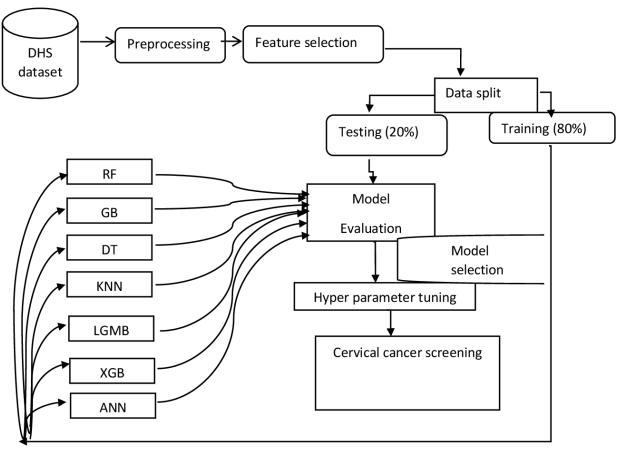
A key strength of machine learning is its ability to handle big datasets while delivering reliable performance metrics. Although this study focuses on sub-Saharan Africa, the application of machine learning in this context allows for the extraction of meaningful patterns from complex health data.

Methods

Data Processing and Management

After deciding on the title by considering the current public health issue, data were extracted from the individual record dataset. Then, the relationships between all variables and the outcome were reviewed based on previous study evidence and expert opinions. The similarity of the selected variables across 4 countries (Ghana, Kenya, Mozambique, and Tanzania) that had recent (2022-2023) Demographic and Health Surveys (DHS) data available was checked. STATA version 17 (StataCorp) and Python 3.10 (Python Software Foundation) were used for data processing and management. To handle missing data, mode imputation and Knearest neighbors imputation were used to maintain the data distribution, ensure statistical consistency, and preserve the integrity of the dataset (Figure 1).

Figure 1. Overview flowchart of data preparation and analysis plan applied. ANN: artificial neural network; DHS: Demographic and Health Surveys; DT: decision tree; GB: gradient boosting; KNN: K-nearest neighbors; LGMB: light gradient boosting machine; RF: random forest; XGB: extreme gradient boosting.



Many machine learning algorithms are sensitive to the scale of features and need normalization before modeling [23]. MinMaxScaler and StandardScaler were applied to normalize or standardize our data, ensuring a consistent scale and improving the performance of modeling tasks. The goal is to minimize performance bias in models that are sensitive to feature scaling, as the absence of standardization can negatively impact model performance.

Feature Selection

Recent studies indicate that irrelevant variables weaken the model's capacity for generalization, raise its overall complexity, and possibly lower a classifier's overall accuracy in a machine learning study [24]. Recursive feature elimination was used to methodically remove irrelevant features, decreasing dimensionality while retaining the most important predictive variables, thereby enhancing the model's efficiency and ability to generalize (Figure 2).

heardabout STI +0.05+0.04wealth status +0.02EDU_level +0.02smartphone Testedfor HIV +0.01+0.01health status1 Sex partner +0.01 marital status -0.01Age first sex Sum of 4 other features +0.02

0.02

0.01

Figure 2. The top 10 features of cervical cancer screening. SHAP: Shapley Additive Explanations.

Splitting the Data

Previous studies indicate that data splitting offers a trustworthy estimation of the model's performance on unseen data and aids in the detection of overfitting [25]. For data splitting, an 80:20 ratio was used for training and testing, respectively, to assess the model's performance and prevent overfitting in machine learning algorithms.

0.00

Handling Imbalanced Data

In machine learning, class imbalance is frequently encountered due to unbalanced data, and this can have a big influence on model accuracy [26].

To improve model generalization and minimize noise and class overlap, we applied the synthetic minority oversampling technique (SMOTE) in combination with Tomek links undersampling. This process resulted in a balanced and cleaner dataset, where the original distribution of 3878 "yes" cases and 30,074 "no" cases was adjusted to 30,074 instances in each class.

Model Selection

After a detailed review of machine learning studies conducted on reproductive health, 7 models were selected

based on simplicity, accuracy, robustness, and computational efficiency, including decision tree, random forest (RF), K-nearest neighbors, extreme gradient boosting, light gradient boosting machine, adaptive boosting, and gradient boosting (GB).

0.04

0.05

Model Training

Both balanced and unbalanced datasets were used to train the selected classifiers after the model was selected. They used 10-fold cross-validation to assess their performance. In order to generate final predictions on test data that had not yet been observed, the best successful predictive model was selected after the comparison and trained using balanced training data.

Hyperparameter Tuning

0.03

mean(|SHAP value|)

A crucial step in maximizing the potential of machine learning models and ensuring their accuracy, efficiency, and robustness is hyperparameter tuning. Hyperparameter tuning by RF is listed in Table 1.

Table 1. Hyperparameters and values for all models.

Classifier	Hyperparameter and value
DT ^a	max_depth=10, criterion=gini, min_samples_split=10, min_sample_leaf=1,
RF ^b	max_depth=20,min_samples_split=10, n_estimators=300
KNN ^c	n_neighbors=5, weights='uniform', metric='minkowski'
ANN ^d	hidden_layer_sizes=(100,), activation='relu', solver='adam', learning_rate='adaptive', max_iter=300
XGB ^e	colsample_bytree=0.8, learning_rate=0.05, max_depth=3, n_estimators=100,
	subsample=1.0
LGBM ^f	learning_rate=0.1, n_estimators=20, num_leaves=31
ADA ^g	n_estimators=300, learning_rate=1.0
GB ^h	n_estimators=300, learning_rate=0.1, max_depth=3, min_samples_split=10

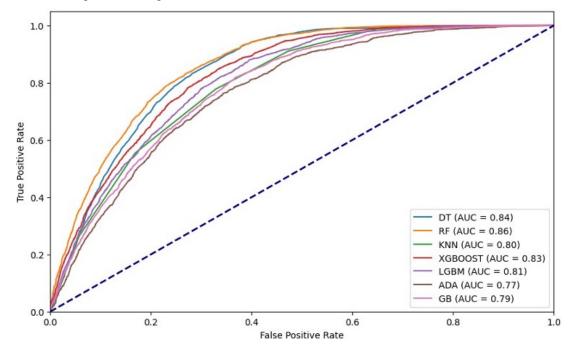
^aDT: decision tree.

Model Evaluation

Model evaluation was performed using confusion matrix, F_1 -score, area under the receiver operating characteristic curve, accuracy, precision, and recall in this study. Ultimately, the RF model outperformed others due to its

effectiveness in handling nonlinear relationships, evaluating feature importance, and minimizing overfitting through the ensemble of multiple decision trees. For results, see Figure 3 and Tables 2 and 3.

Figure 3. ROC curve for model comparison to predict cervical cancer screening. ADA: adaptive boosting; AUC: area under the curve; DT: decision tree; GB: gradient boosting; KNN: K-nearest neighbors; LGMB: light gradient boosting machine; RF: random forest; ROC: receiver operating characteristic; XGB: extreme gradient boosting.



^bRF: random forest.

^cKNN: K-nearest neighbors.

^dANN: artificial neural network.

^eXGB: extreme gradient boosting.

fLGBM: light gradient boosting machine.

^gADA: adaptive boosting. ^hGB: gradient boosting.

Table 2. Model evaluation after synthetic minority oversampling technique.

Model	Accuracy (%)	AUC ^a (%)	F ₁ -score (%)	Recall (%)	Precision (%)
DT ^b	78	85	78	80	76
RF ^c	78	86	79	81	77
KNN ^d	72	80	74	75	73
XGBoost ^e	76	83	76	78	75
LGBM ^f	74	81	75	76	74
ADA ^g	72	77	71	73	69
GB ^h	72	79	72	74	70

^aAUC: area under the curve.

Table 3. Confusion matrix for cervical cancer prediction by random forest accuracy (2802+2494+837+658=6791=78%).

	Predicted positive	Predicted negative
Actual positive	2802 (true positive)	658 (false negative)
Actual negative	837 (false positive)	2494 (true negative)

Ethical Considerations

This study did not require institutional review board approval because it involved secondary analysis of publicly available and deidentified data from the DHS Program. As a result, a consent letter for data access was obtained from a major health and demographic survey program via a web-based request submitted at the DHS Program website [27]. All data used in this study consisted of nonidentifiable information.

Results

Cervical Cancer Screening Distribution by Country

The proportion of women screening rates in Kenya (14%), Mozambique (13%), Tanzania (12%), and Ghana (7%) remains low (Table 4).

Table 4. Cervical cancer screening distribution by country.

Country	Categories	
	Yes, n (%)	No, n (%)
Ghana	530.8 (7)	6767.2 (92)
Kenya	2053 (14)	12,389 (85)
Mozambique	669 (13)	4575.8 (88)
Tanzania	842.6 (12)	6068.1 (88)

Individual Characteristics of Cervical Cancer Screening

The majority of women living in rural (17,987.6/19,690.1, 91%) and urban areas (11,812.7/14,205.8, 83%) have not undergone cervical cancer screening. Among women who own a smartphone, 2170.8 of 11,628.9 (19%) have been screened for cervical cancer, while only 1924.8 of 22,267.0 (9%) of those without a smartphone have been screened.

Women who self-report good health status have a cervical cancer screening rate of 15% (156.9/1,056.6) compared to 11% (2762.9/24,302.6) among those who report poor health status. Women with high wealth status have a cervical cancer screening rate of 18% (2884.8/15,923.9), compared to 5% (604.5/11,398.8) among those in the poorest wealth category.

Only 286.1 of 1875.2 (15%) women who use family planning methods have been screened for cervical cancer. Those who do not consider distance a big problem had a slightly higher screening rate (3226.1/26,048.9, 12%). Women who initiated sexual activity after the age of 30 years had a screening rate of 14% (1932.9/13,527.5), higher than the 11% (2162.7/20,368.2) among those who began before the age of 30 years.

Women reporting multiple sexual partners had a significantly higher screening rate (793.2/4218, 19%) than those who did not (3302.5/29,677.9, 11%). Women who have been tested for HIV have a cervical cancer screening rate of 13% (3958.3/29,777), compared to only 3% (137.3/4118.9) among those who have not been tested. Divorced and

^bDT: decision tree.

cRF: random forest.

^dKNN: K-nearest neighbors.

^eXGBoost: extreme gradient boosting.

fLGBM: light gradient boosting machine.

^gADA: adaptive boosting. ^hGB: gradient boosting.

widowed women show slightly higher screening rates at 15% (676.3/4638.9) and 14% (258.5/1879.7), respectively.

Women with tertiary education have a screening rate of 22% (838.7/3738.8), compared to only 4% (238.6/6572.2) **Table 5.** Individual characteristics of cervical cancer screening.

among those with no education. Women who have heard about sexually transmitted infections (STIs) have a screening rate of 17% (3900.8/23,313.8), compared to only 1.9% (194.8/10,582.1) among those who have not (Table 5).

Variables and categories	Yes, n (%)	No, n (%)
Residence		
Urban	2393.1 (17)	11,812.7 (83)
Rural	1702.5 (9)	17,987.6 (91)
Frequently social media usage		
No	2873.5 (12)	20,629.5 (88)
Yes	1222.2 (12)	9170.8 (88)
Availability of smartphone		
No	1924.8 (9)	20,342.2 (91)
Yes	2170.8 (19)	9458.1 (81)
Self-report of health status		
Good	156.9 (15)	899.7 (85)
Moderate	1175.9 (14)	7360.8 (86)
Poor	2762.9 (11)	21,539.7 (88)
Wealth		
Lowest	604.5 (5)	10,794.3 (94)
Middle	606.3 (9)	5966.9 (91)
Highest	2884.8 (18)	13,039.1 (82)
Family planning method use		
No	3809.5 (12)	28,211.2 (88)
Yes	286.1 (15)	1589.1 (84)
Distance to health facility		
Big problem	869.6 (11)	6977.4 (88)
Not a big problem	3226.1 (12)	22,822.8 (87)
Age at first sexual intercourse (years)		
>30	1932.9 (14)	11,594.6 (85)
<30	2162.7 (11)	18,205.5 (89)
Have multiple sexual partners		
No	3302.5 (11)	26,375.4 (89)
Yes	793.2 (19)	3424.8 (81)
Tested for HIV/AIDS		
No	137.3 (3)	3981.6 (96)
Yes	3958.3 (13)	25,818.7 (86)
Marital status		
Single	254 (13)	1751.7 (87)
Married	2906.7 (12)	22,464.6 (89)
Widowed	258.5 (14)	1621.2 (86)
Divorced	676.3 (15)	3962.6 (85)
Educational status		
No education	238.6 (4)	6333.6 (96)
Primary	1668.7 (12)	12,746.3 (88)
Secondary	1349.6 (15)	7820.2 (85)
Tertiary	838.7 (22)	2900.1 (78)

Variables and categories	Yes, n (%)	No, n (%)	
Heard about STI ^a			
No	194.8 (1.9)	10,387.3 (98)	
Yes	3900.8 (17)	19,413 (83)	

Determinants of Cervical Cancer Screening

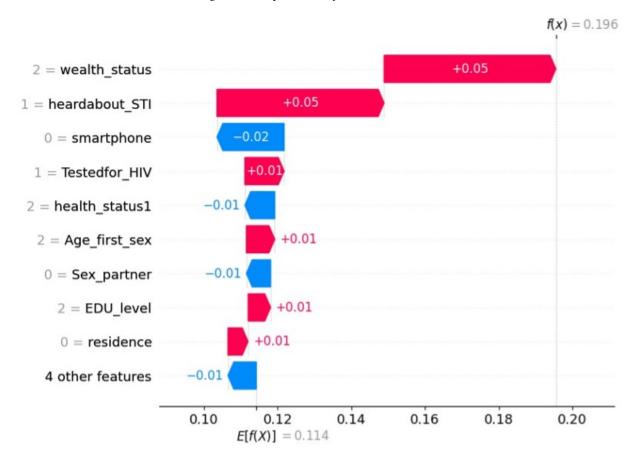
In machine learning, Shapley Additive Explanations (SHAP) values indicate how each feature impacts the output variable. This study examines the impact of the top 10 features on women's cervical cancer screening status. These insights are crucial for specific interventions and policy formulation aimed at enhancing cervical cancer screening practices, as well as reducing morbidity and mortality among women.

In Figure 4, a feature contributing positively to the predicted screening practice is represented by a positive

SHAP value (red), while a feature contributing negatively to the predicted screening outcome is represented by a negative SHAP value (blue).

The best predictive model (RF) showed that wealth status, awareness of STIs, HIV testing, age at first sex, primary education and above, and living in urban areas are significant factors associated with increased cervical cancer screening. However, factors such as not owning a smartphone, having a single sexual partner, and unknown health status are associated with a decrease in cervical cancer screening (Figure 4).

Figure 4. Determinants of cervical cancer screening visualized by a waterfall plot.



Discussion

Principal Findings

In sub-Saharan Africa, cervical cancer is the second most common cause of cancer-related death on the continent and the fourth leading cause of cancer-related death globally [28,29]. Low- and middle-income countries account for nearly 85% of cervical cancer cases and related deaths

[28]. However, according to this study, a small percentage of women were screened for cervical cancer in Kenya (14.22%), Mozambique (12.76%), Tanzania (12.20%), and Ghana (7.27%). This low screening rate contributes to delays in early cancer management and highlights the urgent need for targeted interventions to improve screening behaviors among women.

The predictive evidence from this machine is used to support the realization of the 2030 plan [30]: 90% of cervical diseases identified, 90% of precancerous cases treated, and 90% of women with invasive cancer managed.

The RF is the best-performing model, achieving an accuracy of 78%, an area under the curve of 86%, an F_1 -score of 79%, a recall of 81%, and a precision of 77%. The waterfall plot's SHAP analysis indicates that variables, such as middle and higher wealth status, awareness of STIs, HIV testing, age at first sex, primary education and above, and living in urban areas are significant factors associated with increased cervical cancer screening.

However, not having a smartphone, having a single sexual partner, and being unaware of health status are associated with a decrease in cervical cancer screening.

In this study, the proportion of cervical cancer screening is 13%, which is higher than findings from studies conducted in Cameroon (4%) [2], South India (7.1%) [31], and lower than those reported in Ethiopia (14.79%) [32], and sub-Saharan Africa (19%) [33]. Possible justifications for this difference may include public awareness levels, accessibility of screening services, outreach programs, community health education, socioeconomic status, and geographic barriers. Differences in study populations, sampling methods, and time frames may also contribute to these variations [34].

According to this finding, women with middle and higher wealth status have increased cervical cancer screening rates, which is supported by previous research conducted in the United States, Ethiopia, Uganda, and Cameroon [2,35-37]. One explanation could be that women with middle-class and upper-class incomes have easier access to health care because they can afford the direct and indirect costs of medical care, including transportation and other costs.

This finding shows that women who have information about STIs are more likely to undergo cervical cancer screening. This is supported by previous studies conducted in Southern Ethiopia, Eastern and Southern Africa, Haiti, and Morocco [38-41]. This is because women who have information about STIs are also exposed to information about cervical cancer, including the benefits of early screening, and they tend to have better health-seeking behavior.

According to this SHAP finding, women who do not have smartphones show decreased cervical cancer screening rates. This is supported by previous studies conducted in Japan and England [42,43]. Those who have smartphones can easily access health information, including preventive health care such as screenings, through websites, mobile apps, and social media. Those who do not have smartphones may miss this information and, as a result, may not pay attention to screening [44].

The SHAP finding showed that women who have previously been tested for HIV are more likely to undergo cervical cancer screening. This is consistent with previous studies conducted in Uganda, China, Kenya, and Ethiopia [45-48].

Also, this study indicated that women who did not have information about their health status had decreased cervical cancer screening practices. This is supported by previous studies conducted in Nigeria, Ghana, Australia, India, and Kenya [49-53]. Women who do not have information about their health status are less likely to undergo cervical cancer screening because a lack of awareness reduces their understanding of risk, engagement with health care, and motivation to seek preventive services.

This study showed that women who had their first sexual intercourse before the age of 30 years had an increased cervical screening status. This is in line with studies conducted in East Gojjam Zone, Ethiopia, Africa, and Nigeria [54-56]. A possible explanation for this is that early sexual activity is associated with a higher likelihood of STIs, which in turn leads to more frequent visits to health facilities due to increased awareness of cervical cancer risk.

In addition, women in this age group are more educated and more likely to understand messages from social media and health professionals, including information about cervical cancer risk and the benefits of early screening and treatment [56].

SHAP findings showed that women who have a single sexual partner have decreased cervical cancer screening rates. This finding was supported by studies conducted in Korea, Jordan, Ethiopia, and Uganda [57-60].

Women with a single partner have high confidence that they are free from cervical cancer risk factors because, during health facility visits and through social media information, they understand that women with single sexual partners are at a lower risk of contracting HPV and other STIs, which are the main causes of cervical cancer [61].

This finding also showed that women with primary and higher education levels have increased cervical cancer screening status, which is in line with previous studies conducted in Romania, the Southwest United States, Mayuge, Ghana, and Korea [62-66].

A possible explanation for this could be that educated women have greater health literacy, allowing them to easily understand basic health information and services, including information about cervical cancer. Educated women also tend to have better access to resources such as health care providers, smartphones, and official websites like those of the WHO.

In addition, they may be less influenced by cultural misconceptions or stigma surrounding cervical cancer screening [67].

SHAP findings also showed that women residing in urban areas had increased cervical cancer screening rates. This is in line with previous studies conducted in Nepal, Nigeria, China, and Ghana [55,68-70]. Women residing in urban areas have improved access to health information and services through social media, public health campaigns, and proximity to health care facilities and professionals. Consequently,

they tend to have a better understanding of cervical cancer compared to their rural counterparts [71].

Strengths and Limitations

One of the study's strengths is that it makes use of a large and varied dataset, which improves the findings' applicability to a larger geographic area. Furthermore, complex patterns and relationships that traditional statistics can miss might be found through the use of machine learning models, and the limitations of this include issues that can impact prediction accuracy, such as self-reported information, recall bias, and missing data. Furthermore, it is challenging to determine causal links due to the cross-sectional nature of data, and the findings based on machine learning outputs should be interpreted as associations rather than causations.

Conclusions

Wealth status, awareness of STIs, HIV testing exposure, age at first sexual intercourse, educational level, residency, smartphone ownership, having a single sexual partner, and previous health status are predictors of cervical cancer screening.

Improving education and awareness, expanding access to screening, especially in rural areas, leveraging both digital health and community-based outreach, integrating screening with other health services, and addressing socioeconomic barriers are recommended strategies to increase cervical cancer screening rates in sub-Saharan Africa.

Acknowledgments

The authors would like to express their gratitude for the valuable time and expertise provided by everyone involved in this study. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets analyzed in this study are available in the public domain through the Measure DHS website [72].

Authors' Contributions

NDB, GGA, and MNA developed the concept for the study. AM, LDB, MCA, HWA, DGA, TGA, and ATB reviewed the literature. NDB, GGA, MNA, and HWA conducted the data analysis. NDB, GGA, MNA, HWA, and DGA discussed the findings.

Conflicts of Interest

None declared.

References

- 1. Cervical cancer. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer [Accessed 2025-08-18]
- 2. Okyere J, Duodu PA, Aduse-Poku L, Agbadi P, Nutor JJ. Cervical cancer screening prevalence and its correlates in Cameroon: secondary data analysis of the 2018 demographic and health surveys. BMC Public Health. Jun 5, 2021;21(1):1071. [doi: 10.1186/s12889-021-11024-z] [Medline: 34090372]
- 3. Amorha KC, Ozota GO, Ndunwere MGO, Anyaji UL, Egbo OF, Ogugofor OA. Knowledge, attitudes, and practices of adolescent girls regarding cervical cancer: a cross-sectional study in Enugu State, Nigeria. Pan Afr Med J. 2024;47:17. [doi: 10.11604/pamj.2024.47.17.41087] [Medline: 38524107]
- 4. Olivieri DJ, Eastment MC, Mugisha N, Menon MP. Correlates of cervical cancer awareness among women aged 30-49 in five sub-Saharan African nations: evidence from the Demographic and Health Survey (DHS)-2017-2023. PLOS Glob Public Health. 2025;5(5):e0003344. [doi: 10.1371/journal.pgph.0003344] [Medline: 40333666]
- 5. Mantula F, Toefy Y, Sewram V. Barriers to cervical cancer screening in Africa: a systematic review. BMC Public Health. Feb 20, 2024;24(1):525. [doi: 10.1186/s12889-024-17842-1] [Medline: 38378542]
- 6. Uddin KMM, Al Mamun A, Chakrabarti A, Mostafiz R, Dey SK. An ensemble machine learning-based approach to predict cervical cancer using hybrid feature selection. Neuroscience Informatics. Sep 2024;4(3):100169. [doi: 10.1016/j.neuri.2024.100169]
- 7. Saaka SA, Hambali MG. Factors associated with cervical cancer screening among women of reproductive age in Ghana. BMC Womens Health. Sep 18, 2024;24(1):519. [doi: 10.1186/s12905-024-03367-7] [Medline: 39289681]
- 8. Shakil R, Islam S, Akter B. A precise machine learning model: detecting cervical cancer using feature selection and explainable AI. J Pathol Inform. Dec 2024;15:100398. [doi: 10.1016/j.jpi.2024.100398] [Medline: 39494050]
- 9. Ong SK, Abe SK, Thilagaratnam S, et al. Towards elimination of cervical cancer human papillomavirus (HPV) vaccination and cervical cancer screening in Asian National Cancer Centers Alliance (ANCCA) member countries. Lancet Reg Health West Pac. Oct 2023;39:100860. [doi: 10.1016/j.lanwpc.2023.100860] [Medline: 37576906]

- 10. Advancing the cervical cancer elimination agenda in the African region. WHO Regional Office for Africa; Nov 2024. URL: https://www.afro.who.int/media-centre/statements-commentaries/advancing-cervical-cancer-elimination-agenda-african-region [Accessed 2025-09-10]
- 11. Human Papillomavirus (HPV) Infection. International Agency for Research on Cancer; 2007. URL: https://www.ncbi.nlm.nih.gov/books/NBK321770/ [Accessed 2025-09-10]
- 12. Beyene T, Akibu M, Bekele H, Seyoum W. Risk factors for precancerous cervical lesion among women screened for cervical cancer in south Ethiopia: unmatched case-control study. PLoS One. 2021;16(7):e0254663. [doi: 10.1371/journal.pone.0254663] [Medline: 34265004]
- 13. Holcakova J, Bartosik M, Anton M, et al. New trends in the detection of gynecological precancerous lesions and early-stage cancers. Cancers (Basel). Dec 17, 2021;13(24):6339. [doi: 10.3390/cancers13246339] [Medline: 34944963]
- 14. Yan T, Wong PK, Qin YY. Deep learning for diagnosis of precancerous lesions in upper gastrointestinal endoscopy: a review. World J Gastroenterol. May 28, 2021;27(20):2531-2544. [doi: 10.3748/wjg.v27.i20.2531] [Medline: 34092974]
- 15. Yang L, Boily MC, Rönn MM, et al. Regional and country-level trends in cervical cancer screening coverage in sub-Saharan Africa: a systematic analysis of population-based surveys (2000-2020). PLoS Med. Jan 2023;20(1):e1004143. [doi: 10.1371/journal.pmed.1004143] [Medline: 36634119]
- 16. Adewumi K, Nishimura H, Oketch SY, Adsul P, Huchko M. Barriers and facilitators to cervical cancer screening in Western Kenya: a qualitative study. J Cancer Educ. Aug 2022;37(4):1122-1128. [doi: 10.1007/s13187-020-01928-6] [Medline: 33411253]
- 17. Tshabalala G, Blanchard C, Mmoledi K, et al. A qualitative study to explore healthcare providers' perspectives on barriers and enablers to early detection of breast and cervical cancers among women attending primary healthcare clinics in Johannesburg, South Africa. PLOS Glob Public Health. 2023;3(5):e0001826. [doi: 10.1371/journal.pgph.0001826] [Medline: 37159437]
- 18. Tenkir L, Mamuye A, Jemebere W, Yeheyis T. The magnitude of precancerous cervical lesions and its associated factors among women screened for cervical cancer at a referral center in Southern Ethiopia, 2021: a cross-sectional study. Front Glob Womens Health. 2023;4:1187916. [doi: 10.3389/fgwh.2023.1187916] [Medline: 37664421]
- 19. Ampofo AG, Boyes AW, Asibey SO, Oldmeadow C, Mackenzie LJ. Prevalence and correlates of modifiable risk factors for cervical cancer and HPV infection among senior high school students in Ghana: a latent class analysis. BMC Public Health. Feb 15, 2023;23(1):340. [doi: 10.1186/s12889-022-14908-w] [Medline: 36793003]
- 20. El-Zein M, Ramanakumar AV, Naud P, et al. Determinants of acquisition and clearance of human papillomavirus infection in previously unexposed young women. Sex Transm Dis. Oct 2019;46(10):663-669. [doi: 10.1097/OLQ. 000000000001053] [Medline: 31464859]
- 21. Guo C, Zhan B, Li MY, Yue L, Zhang C. Association between oral contraceptives and cervical cancer: a retrospective case-control study based on the National Health and Nutrition Examination Survey. Front Pharmacol. 2024;15:1400667. [doi: 10.3389/fphar.2024.1400667] [Medline: 39086392]
- 22. Torres-Poveda K, Ruiz-Fraga I, Madrid-Marina V, Chavez M, Richardson V. High risk HPV infection prevalence and associated cofactors: a population-based study in female ISSSTE beneficiaries attending the HPV screening and early detection of cervical cancer program. BMC Cancer. Dec 10, 2019;19(1):1205. [doi: 10.1186/s12885-019-6388-4] [Medline: 31823749]
- 23. Feature scaling in machine learning: standardization, MinMaxScaling and more. Train In Data. URL: https://www.blog.trainindata.com/feature-scaling-in-machine-learning/#:~:text=Many%20machine%20learning%20algorithms%20are%20sensitive [Accessed 2025-08-18]
- 24. Vyas A, Raman S, Sen S, et al. Machine learning-based diagnosis and ranking of risk factors for diabetic retinopathy in population-based studies from South India. Diagnostics (Basel). Jun 16, 2023;13(12):2084. [doi: 10.3390/diagnostics13122084] [Medline: 37370980]
- Muraina I. Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. Presented at: 7th International Mardin Artuklu scientific research conference; Dec 10-12, 2021; Mardin, Turkey. URL: https://tinyurl.com/2j7j9ezm [Accessed 2025-09-04]
- 26. 10 techniques to handle imbalanced classes in machine learning. Analytics Vidhya. URL: https://www.analyticsvidhya.com/articles/class-imbalance-in-machine-learning [Accessed 2025-08-18]
- 27. The DHS Program. URL: https://dhsprogram.com/ [Accessed 2025-08-18]
- 28. Cervical cancer. World Health Organization African region. URL: https://www.afro.who.int/health-topics/cervical-cancer [Accessed 2025-08-18]
- 29. Shocking statistics: cervical cancer claims 70% of cases in Africa, experts call for action. Top Africa news. URL: https://www.topafricanews.com/2025/04/24 [Accessed 2025-08-18]

- 30. Continental consultative meeting report: accelerating the plan to eliminate cervical cancer in Africa by 2030. AfricaCDC; URL: https://africacdc.org/download/continental-consultative-meeting-report-accelerating-the-plan-to-eliminate-cervical-cancer-in-africa-by-2030/ [Accessed 2025-09-04]
- 31. Reichheld A, Mukherjee PK, Rahman SM, David KV, Pricilla RA. Prevalence of cervical cancer screening and awareness among women in an urban community in South India-a cross sectional study. Ann Glob Health. Mar 16, 2020;86(1):30. [doi: 10.5334/aogh.2735] [Medline: 32211300]
- 32. Desta M, Getaneh T, Yeserah B, et al. Cervical cancer screening utilization and predictors among eligible women in Ethiopia: a systematic review and meta-analysis. PLoS One. 2021;16(11):e0259339. [doi: 10.1371/journal.pone. 0259339] [Medline: 34735507]
- 33. Ba DM, Ssentongo P, Musa J, et al. Prevalence and determinants of cervical cancer screening in five sub-Saharan African countries: a population-based study. Cancer Epidemiol. Jun 2021;72:101930. [doi: 10.1016/j.canep.2021. 101930] [Medline: 33756434]
- 34. Asgedom YS, Kassie GA, Habte A, Ketema DB, Aragaw FM. Socioeconomic inequality in cervical cancer screening uptake among women in sub-Saharan Africa: a decomposition analysis of Demographic and Health Survey data. BMJ Open. Dec 10, 2024;14(12):e088753. [doi: 10.1136/bmjopen-2024-088753] [Medline: 39658286]
- 35. Isabirye A, Mbonye MK, Kwagala B. Predictors of cervical cancer screening uptake in two districts of Central Uganda. PLoS ONE. 2020;15(12):e0243281. [doi: 10.1371/journal.pone.0243281] [Medline: 33270792]
- 36. Miles-Richardson S, Allen S, Claridy MD, Booker EA, Gerbi G. Factors associated with self-reported cervical cancer screening among women aged 18 years and older in the United States. J Community Health. Feb 2017;42(1):72-77. [doi: 10.1007/s10900-016-0231-5] [Medline: 27496176]
- 37. Misgun T, Demissie DB. Knowledge, practice of cervical cancer screening and associated factors among women police members of Addis Ababa police commission Ethiopia. BMC Cancer. Oct 10, 2023;23(1):961. [doi: 10.1186/s12885-023-11478-x] [Medline: 37817082]
- 38. Amado G, Weldegebreal F, Birhanu S, Dessie Y. Cervical cancer screening practices and its associated factors among females of reproductive age in Durame town, Southern Ethiopia. PLoS ONE. 2022;17(12):e0279870. [doi: 10.1371/journal.pone.0279870] [Medline: 36584208]
- 39. Bendahhou K, Serhier Z, Diouny S, et al. Women's knowledge and attitudes towards cervical cancer screening in Morocco. Cureus. Apr 2023;15(4):e37989. [doi: 10.7759/cureus.37989] [Medline: 37223139]
- 40. Macleod CI, Reynolds JH. Human papilloma virus infection and cervical cancer among women who sell sex in Eastern and Southern Africa: a scoping review. Womens Health (Lond). 2021;17:17455065211058349. [doi: 10.1177/17455065211058349] [Medline: 34775848]
- 41. McCarthy SH, Walmer KA, Boggan JC, et al. Awareness of cervical cancer causes and predeterminants of likelihood to screen among women in Haiti. J Low Genit Tract Dis. Jan 2017;21(1):37-41. [doi: 10.1097/LGT.000000000000000281] [Medline: 27906806]
- 42. Groves S, Brooks J. What do young women below national screening age in England think about cervical cancer and cervical screening? A qualitative study. J Clin Nurs. Jun 2022;31(11-12):1588-1597. [doi: 10.1111/jocn.16012] [Medline: 34418192]
- 43. Irino S, Ose H, Takata N, Kamoshida S, Ohsaki H. Barriers to undergoing cervical cancer screening among health sciences university students in Japan: a cross-sectional study. Nurs Health Sci. Sep 2023;25(3):466-473. [doi: 10.1111/nhs.13043] [Medline: 37581364]
- 44. Ghahramani F, Wang J. Impact of smartphones on quality of life: a health information behavior perspective. Inf Syst Front. Dec 2020;22(6):1275-1290. [doi: 10.1007/s10796-019-09931-z]
- 45. Choi Y, Ibrahim S, Park LP, Cohen CR, Bukusi EA, Huchko MJ. Uptake and correlates of cervical cancer screening among women attending a community-based multi-disease health campaign in Kenya. BMC Womens Health. Apr 18, 2022;22(1):122. [doi: 10.1186/s12905-022-01702-4] [Medline: 35436908]
- 46. Emru K, Abebaw TA, Abera A. Role of awareness on cervical cancer screening uptake among HIV positive women in Addis Ababa, Ethiopia: a cross-sectional study. Womens Health (Lond). 2021;17:17455065211017041. [doi: 10.1177/17455065211017041] [Medline: 34096400]
- 47. Lin S, Chen WT, Gu C, Cheng HL, Wang H, Tang S. Knowledge, perception of HIV symptom severity and cervical cancer screening behaviour among women living with HIV in China. Eur J Cancer Care (Engl). Mar 2022;31(2):e13542. [doi: 10.1111/ecc.13542] [Medline: 34961999]
- 48. Ninsiima M, Nyabigambo A, Kagaayi J. Acceptability of integration of cervical cancer screening into routine HIV care, associated factors and perceptions among HIV-infected women: a mixed methods study at Mbarara Regional Referral Hospital, Uganda. BMC Health Serv Res. Apr 3, 2023;23(1):333. [doi: 10.1186/s12913-023-09326-6] [Medline: 37013535]

- 49. Brown RF, Muller TR, Olsen A. Australian women's cervical cancer screening attendance as a function of screening barriers and facilitators. Soc Sci Med. Jan 2019;220:396-402. [doi: 10.1016/j.socscimed.2018.11.038] [Medline: 30529797]
- 50. Eastment MC, Wanje G, Richardson BA, et al. A cross-sectional study of the prevalence, barriers, and facilitators of cervical cancer screening in family planning clinics in Mombasa County, Kenya. BMC Health Serv Res. Dec 23, 2022;22(1):1577. [doi: 10.1186/s12913-022-08984-2] [Medline: 36564740]
- 51. Ghosh S, Mallya SD, Shetty RS, et al. Knowledge, attitude and practices towards cervical cancer and its screening among women from tribal population: a community-based study from Southern India. J Racial Ethn Health Disparities. Feb 2021;8(1):88-93. [doi: 10.1007/s40615-020-00760-4] [Medline: 32333376]
- 52. Rimande-Joel R, Ekenedo GO. Knowledge, belief and practice of cervical cancer screening and prevention among women of Taraba, North-East Nigeria. Asian Pac J Cancer Prev. Nov 1, 2019;20(11):3291-3298. [doi: 10.31557/APJCP. 2019.20.11.3291] [Medline: 31759351]
- 53. Stuart A, Obiri-Yeboah D, Adu-Sarkodie Y, Hayfron-Benjamin A, Akorsu AD, Mayaud P. Knowledge and experience of a cohort of HIV-positive and HIV-negative Ghanaian women after undergoing human papillomavirus and cervical cancer screening. BMC Womens Health. Oct 23, 2019;19(1):123. [doi: 10.1186/s12905-019-0818-y] [Medline: 31647013]
- 54. Mekonnen AG, Mittiku YM. Early-onset of sexual activity as a potential risk of cervical cancer in Africa: a review of literature. PLOS Glob Public Health. 2023;3(3):e0000941. [doi: 10.1371/journal.pgph.0000941] [Medline: 36962975]
- 55. Olubodun T, Ogundele OO, Salisu ZA, Odusolu YO, Caleb-Ugwuowo UU. Cervical cancer awareness and risk factors among women residing in an urban slum in Lagos, Southwest Nigeria. Afr Health Sci. Sep 2023;23(3):269-279. [doi: 10.4314/ahs.v23i3.33] [Medline: 38357134]
- 56. Tesfaw K, Kindie W, Mulatu K, Bogale EK. Utilisation of cervical cancer screening and factors associated with screening utilisation among women aged 30-49 years in Mertule Mariam Town, East Gojjam Zone, Ethiopia, in 2021: a cross-sectional survey. BMJ Open. Nov 22, 2022;12(11):e067229. [doi: 10.1136/bmjopen-2022-067229] [Medline: 36414295]
- 57. Isabirye A. Individual and intimate-partner factors associated with cervical cancer screening in Central Uganda. PLoS One. 2022;17(9):e0274602. [doi: 10.1371/journal.pone.0274602] [Medline: 36108074]
- 58. Kim S, Lee SY, Choi-Kwon S. Cervical cancer screening and human papillomavirus vaccination among Korean sexual minority women by sex of their sexual partners. Int J Environ Res Public Health. Nov 30, 2020;17(23):8924. [doi: 10.3390/ijerph17238924] [Medline: 33266281]
- 59. Teame H, Gebremariam L, Kahsay T, Berhe K, Gebreheat G, Gebremariam G. Factors affecting utilization of cervical cancer screening services among women attending public hospitals in Tigray region, Ethiopia, 2018; case control study. PLoS One. 2019;14(3):e0213546. [doi: 10.1371/journal.pone.0213546] [Medline: 30870497]
- 60. Urquhart G, Maclennan SJ, Guntupalli AM. Is there an association between intimate partner violence and the prevalence of cervical cancer screening in Jordan? PLoS One. 2023;18(8):e0290678. [doi: 10.1371/journal.pone.0290678] [Medline: 37651440]
- 61. CervixCheck information for healthcare providers. CancerCare Manitoba. URL: https://www.cancercare.mb.ca/screening/hcp/cervix [Accessed 2025-09-04]
- 62. Cartwright K, Kosich M, Gonya M, et al. Cervical cancer knowledge and screening patterns in Zuni Pueblo Women in the Southwest United States. J Cancer Educ. Oct 2023;38(5):1531-1538. [doi: 10.1007/s13187-023-02295-8] [Medline: 370461421
- 63. Covaliu BF, Forray AI, Tomic M, et al. Understanding cervical cancer screening attendance: barriers and facilitators in a representative population survey. Cancers (Basel). Feb 19, 2025;17(4):706. [doi: 10.3390/cancers17040706] [Medline: 40002299]
- 64. Fang CY, Ma GX, Handorf EA, et al. Addressing multilevel barriers to cervical cancer screening in Korean American women: a randomized trial of a community-based intervention. Cancer. May 15, 2017;123(6):1018-1026. [doi: 10.1002/cncr.30391] [Medline: 27869293]
- 65. Nakisige C, Trawin J, Mitchell-Foster S, et al. Integrated cervical cancer screening in Mayuge District Uganda (ASPIRE Mayuge): a pragmatic sequential cluster randomized trial protocol. BMC Public Health. Jan 31, 2020;20(1):142. [doi: 10.1186/s12889-020-8216-9] [Medline: 32005202]
- 66. Williams MS, Kenu E, Adanu A, et al. Awareness and beliefs about cervical cancer, the HPV vaccine, and cervical cancer screening among Ghanaian women with diverse education levels. J Cancer Educ. Oct 2019;34(5):897-903. [doi: 10.1007/s13187-018-1392-y] [Medline: 29974412]
- 67. Rosyda R, Santoso B, Yunitasari E. Is an educational level affect women's participation on cervical cancer screening? a systematic review. Presented at: The 9th International Nursing Conference; Apr 7-8, 2018; Surabaya East Java, Indonesia. [doi: 10.5220/0008327204820488]

- 68. Adzigbli LA, Aboagye RG, Adeleye K, Osborne A, Ahinkorah BO. Cervical cancer screening uptake and its predictors among women aged 30-49 in Ghana: providing evidence to support the World Health Organization's cervical cancer elimination initiative. BMC Infect Dis. Feb 21, 2025;25(1):246. [doi: 10.1186/s12879-025-10485-6] [Medline: 39984848]
- 69. Paneru B, Karmacharya A, Bharati A, et al. Association between cancer stigma and cervical cancer screening uptake among women of Dhulikhel and Banepa, Nepal. PLoS ONE. 2023;18(5):e0285771. [doi: 10.1371/journal.pone.0285771] [Medline: 37200234]
- 70. Zhang M, Zhong Y, Zhao Z, et al. Cervical cancer screening rates among Chinese women China, 2015. China CDC Wkly. Jun 26, 2020;2(26):481-486. [doi: 10.46234/ccdcw2020.128] [Medline: 34594684]
- 71. Zhetpisbayeva I, Kassymbekova F, Sarmuldayeva S, Semenova Y, Glushkova N. Cervical cancer prevention in rural areas. Ann Glob Health. 2023;89(1):75. [doi: 10.5334/aogh.4133] [Medline: 37928103]
- 72. Measuredhs. URL: https://www.measuredhs.com/ [Accessed 2025-08-19]

Abbreviations

DHS: Demographic and Health Survey

LR: logistic regression **RF:** random forest

SHAP: Shapley Additive Explanations **STI:** sexually transmitted infection **WHO:** World Health Organization

Edited by John Tayu Lee; peer-reviewed by Ayenew Gebeyew, L Raymond Guo; submitted 23.01.2025; final revised version received 28.05.2025; accepted 06.06.2025; published 17.09.2025

Please cite as:

Baykemagn ND, Aweke MN, Mesfin A, Baffa LD, Agimas MC, Abuhay HW, Adugna DG, Alemu TG, Bicha AT, Getu Alemu G

Identifying Predictors of Cervical Cancer Screening Uptake in Sub-Saharan Africa Using Machine Learning: Cross-Sectional Study

JMIR Public Health Surveill 2025;11:e71677 URL: https://publichealth.jmir.org/2025/1/e71677

doi: 10.2196/71677

© Nebebe Demis Baykemagn, Mekuriaw Nibret Aweke, Amare Mesfin, Lemlem Daniel Baffa, Muluken Chanie Agimas, Habtamu Wagnew Abuhay, Dagnew Getnet Adugna, Tewodros Getaneh Alemu, Alemu Teshale Bicha, Gebrie Getu Alemu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.09.2025. This is an openaccess article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.