Communication in Physical Sciences, 2026, 13(7):1040-1055

Socio-Demographic Determinants of HIV Infection Status among Patients Attending Public Hospitals in Akwa Ibom North-West Senatorial District, Nigeria

Unyime Patrick Udoudo*, Eduma E. Essien and Emmanuel J. Ekpenyong

Received: 14 April 2026/Accepted: 24 June 2026 /Published: 08 July 2026

https://dx.doi.org/10.4314/cps.v13i7.2

Abstract: *This study investigated the influence of socio-demographic characteristics on the HIV infection status of patients attending public hospitals in Akwa Ibom North-West Senatorial District, Nigeria. Data were obtained from patients attending four randomly selected public hospitals between 2021 and 2025. The study covered eight public hospitals across the senatorial district, from which four hospitals were selected using simple random sampling. A total of 2,611 patient records were analysed, comprising 1,581 (60.6%) HIV-negative and 1,030 (39.4%) HIV-positive patients. The explanatory variables included family history of HIV infection, sex, marital status, educational level, employment status, place of residence, age, CD4 count, and growth phase, while HIV infection status served as the binary response variable. Descriptive statistics and multiple binary logistic regression were used for data analysis.The results showed that family history of HIV infection was the strongest predictor of HIV-positive status (AOR = 93.31, 95% CI: 35.28–246.80, $p < 0.001$). Male patients had significantly higher odds of HIV infection than females (AOR = 1.774, 95% CI: 1.491–2.110, $p < 0.001$). Compared with the reference categories, patients with primary (AOR = 1.748, $p = 0.006$) and secondary education (AOR = 2.089, $p = 0.001$), government-employed patients (AOR = 1.260, $p = 0.036$), and patients aged 11–20 years (AOR = 2.187, $p = 0.038$), 21–30 years (AOR = 3.134, $p = 0.003$), 31–40 years (AOR = 3.495, $p = 0.001$), 41–50 years (AOR = 4.625, $p < 0.001$), and 51–60 years (AOR = 4.034, $p = 0.001$) had significantly higher odds of testing HIV-positive. Conversely, urban residence was associated with significantly lower odds of HIV infection than rural residence (AOR = 0.818, 95% CI: 0.688–0.973, $p = 0.023$). The logistic regression model demonstrated an adequate fit to the data (Hosmer–Lemeshow $\chi^2 = 10.661$, df = 8, $p = 0.222$) and an overall classification accuracy of 68.9%. The study concludes that family history of HIV infection, sex, marital status, educational level, employment status, residence, and age are significant determinants of HIV infection status among patients attending public hospitals in Akwa Ibom North-West Senatorial District. The findings underscore the need for targeted HIV prevention, screening, and public health interventions focusing on socio-demographic groups at greater risk of HIV infection.*

Keywords*:* *HIV infection status, socio-demographic characteristics, logistic regression, adjusted odds ratio (AOR), Akwa Ibom North-West Senatorial District, patients, hospitals.*

Unyime Patrick Udoudo
Department of Statistics, Akwa Ibom State Polytechnic, Ikot Osurua, Akwa Ibom State, Nigeria.
Email: udoudogeno@gmail.com

Eduma E. Essien
Federal Polytechnic, Ukana, Akwa Ibom State, Nigeria.
Email: edumaessien@gmail.com

Available @ https://doi.org/10.5281/zenodo.21294313

Emmanuel J. Ekpenyong
Department of Mathematics and Statistics, Federal University of Technology, Ikot Abasi, Nigeria.
Email: emmanuel.ekpenyong@futia.edu.ng
https://orcid.org/0000-0003-4112-3712

1.0 Introduction

Akwa Ibom State has consistently recorded one of the highest HIV prevalence rates in Nigeria, making it a priority area for HIV prevention and control interventions (Abogoh *et al*., 2025; NACA, 2024; Utip et al., 2025) Nigeria, with over two million people estimated to be living with HIV, faces a substantial public health challenge (NACA Spectrum Estimates, 2024; Nastiti *et al*., 2024). Globally, HIV remains one of the leading infectious diseases, affecting millions of people despite considerable progress in prevention, diagnosis, and access to antiretroviral therapy. Understanding the factors associated with HIV infection remains essential for reducing new infections and improving treatment outcomes. Despite significant progress in HIV prevention and treatment, including expanded access to antiretroviral therapy (ART), Akwa Ibom State continues to record the highest HIV prevalence in Nigeria, with an estimated prevalence of 5.5%, indicating the need for sustained epidemiological investigations into factors influencing HIV transmission, (Tribune Newspaper, 2025). The persistently high prevalence of HIV in Akwa Ibom State suggests that several socio-demographic, behavioural, and environmental factors may contribute to continued HIV transmission. Identifying these determinants is essential for developing targeted interventions capable of reducing HIV incidence and improving public health outcomes. Several studies have investigated the influence of socio-demographic characteristics on HIV infection, treatment outcomes, and survival across different countries and populations. Tekle (2019) examined socio-economic, demographic, and health factors influencing survival versus death among HIV-positive patients under ART follow-up at Wolaita Sodo Referral Hospital in Ethiopia. Binary logistic regression identified several significant predictors of mortality. Socio-economic and demographic factors (age, sex, marital status, education, residence), alongside clinical factors (CD4 count, WHO stage, opportunistic infections), were significant predictors of survival among HIV-positive patients under ART follow-up. The study concludes that early diagnosis, timely ART initiation, and targeted support for vulnerable groups (older patients, rural residents, widowed/divorced individuals, and those with low education) are essential to improve survival outcomes. Strengthening community-based interventions, health education, and equitable access to ART services is critical in reducing HIV-related mortality in Ethiopia. Tugga *et al*. (2026) work on estimation of binary logistic regression using three Links Function (Logit, Probit, and Complementary Log Log) in assessing the factors that influence HIV. They found that there is a consistent decline in HIV odds across the study years, significantly higher odds among females, and substantially increased odds among adults aged 30–49 years and those above 50 years, and concluded that age, sex, and years are significant predictors of HIV infection. Lawan *et al*. (2016), in their work found that awareness of HIV status and risks were significantly associated with age, education, and risk behaviors. Despite high awareness levels, risky transmission behaviors remained prevalent, leading to the conclusion that awareness alone is insufficient. The work considered sample size of 400 HIV-positive adolescents in Kano, analyzed using SPSS with logistic regression and chi-square tests. Howard (2024) revealed that enhancement in predictive accuracy for a logistic regression is possible by means of incorporating PCA and K-Medians with robust centers. Model 5 was found to be the best predictor of HIV/AIDS status of a patient. It is an integration of both robust principal component analysis and K-Medians clustering into a binary logistic regression model. Gorfu *et al* (2026). show that the prevalence of HIV was marginally higher in women (3.94%) than in men (3.15%) and in urban dwellers (3.8%) than in rural ones (3.1%). Participants who were divorced or widowed had significantly higher odds of HIV infection than those who were single after controlling for potential confounders (COR = 4.62% and COR = 4.43%, respectively). The odds of being HIV-positive were 10.57 times higher for participants who reported having three or more lifetime sexual partners than for those who reported having no lifetime sexual partners (COR = 10.57%). The study demonstrated a notable prevalence of HIV infection in the study area, with significant associations observed with socio-demographic and behavioral factors. Overall, previous studies consistently demonstrate that age, sex, marital status, education, residence, and clinical characteristics significantly influence HIV infection or treatment outcomes, although the strength and direction of these associations vary across different populations and geographical settings. Mendez *et al*. (2023) discovered using the 2017–2019 National Survey of Family Growth data, when seven survey-weighted multinomial/binary logistic regression models were analyzed in men ages 15–49 years old (N = 4260) and found that specific health care provider (HCP) discussions significantly increased the odds of HIV and STI screening among men. Noted also that sexual partners, condom use, sexual orientation, and discussing HIV/AIDS were the strongest predictors of screening uptake and conclude that targeted provider-patient conversations can meaningfully promote HIV/STI testing. However, many of these studies focused on behavioural or clinical determinants, while relatively few comprehensively assessed multiple socio-demographic characteristics among patients attending public hospitals using routinely collected hospital data Gaughan *et al*. (2026) discovered that unhoused patients in Tijuana had an HIV prevalence of 10%, nearly 50 times higher than Mexico's national average (0.19%). Binary logistic regression showed that women, people who use drugs, and those with another STI were significantly more likely to test HIV-positive. The conclusion emphasizes the urgent need for expanded HIV rapid testing and linkage to care, including pre-exposure prophylaxis (PrEP), in the vulnerable population. Iseh, *et al*. (2022) in Juxtaposing Vertically Transmitted Infections (VTIs) and the Spread of HIV/AIDS in a Typically Infection Prevalent Region in Nigeria, found that vertically transmitted infections (VTIs) such as syphilis, hepatitis B, and hepatitis C significantly increased the likelihood of HIV spread in Nigeria. Binary logistic regression showed that co-infections were strong predictors of HIV positivity, with maternal-to-child transmission pathways amplifying risk. The conclusion emphasizes that controlling VTIs is essential to reducing HIV prevalence in infection-prone regions.

Although previous studies have identified several socio-demographic determinants of HIV infection, most were conducted outside Akwa Ibom State, examined relatively few explanatory variables, or focused on specific population groups rather than hospital-based patients. Furthermore, few studies have investigated HIV infection status among patients attending public hospitals across multiple Local Government Areas within a senatorial district. Consequently, there is limited localized evidence on the combined effects of socio-demographic characteristics on HIV infection status in Akwa Ibom North-West Senatorial District. Therefore, the aim of this study was to investigate the impacts of socio-demographic characteristics on HIV infection status among patients attending selected public hospitals in Akwa Ibom North-West Senatorial District using multiple binary logistic regression analysis. The findings of this study will provide evidence for healthcare providers, public health practitioners, and policymakers regarding the socio-demographic determinants of HIV infection within the study area. The results are expected to support targeted HIV prevention strategies, improve resource allocation, strengthen evidence-based public health planning, and contribute to the growing body of literature on HIV epidemiology in Nigeria.

2.0 Materials and Methods

2.1 Data collection

This study utilized secondary data extracted from the medical records of patients who attended public hospitals for medical attention across the ten (10) Local Government Areas (LGAs) of the Akwa Ibom North-West Senatorial District. These LGAs include Ikot Ekpene, Ikono, Ini, Essien Udim, Obot Akara, Abak, Etim Ekpo, Ukanafun, Oruk Anam, and Ika. Among these ten, Obot Akara and Ika do not have public general hospitals; therefore, eight (8) hospitals comprised the eligible institutional population.

A simple random sampling technique was used to select four (4) hospitals from the eight available. Information regarding patients' HIV infection status and socio-demographic characteristics was extracted for the period spanning 2021 to 2025. The socio-demographic variables collected include age, sex, residence, marital status, educational level, employment status, family history of HIV infection, CD4 count, and growth phases. The dependent variable is dichotomous, representing the HIV infection status of the patient (positive or negative). A total sample size of 2,611 patient records was observed.

Because the exact target population size within the timeframe was unknown, a single-stage cluster sampling method was adopted, where each hospital was treated as a cluster. A simple random sample of size \$n\$ clusters was drawn without replacement. The sample size calculation incorporated the known 5.5% HIV prevalence rate of Akwa Ibom State. To estimate the baseline proxy for a single senatorial district, the state prevalence was adjusted proportionally across the three senatorial districts (5.5% / 3 = 1.83%).

In this study, the population size is unknown; therefore, the sampling method adopted is the single stage cluster sampling, in which each hospital is regarded as a cluster. A simple random sample of size n from these clusters is drawn without replacement. The sample size is determined using the formula

$$n = \frac{Np(1-p)}{(N-1)\frac{d^2}{Z_{\frac{\alpha}{2}}} + p(1-p)} = \frac{Npq}{(N-1)d^2 Z_{\frac{\alpha}{2}}^{-2} + pq} \qquad (1)$$

where, $N =$ population size, $d =$ margin of error taken to be 0.1. p is estimated from the prevalence rate of 5.5% of HIV infection in Akwa Ibom State, for the three (3) senatorial districts. To have an estimate of p for one senatorial district, 5.5% is divided by 3.

$$p = \frac{5.5\%}{3} = 1.033\% = 0.0183; \; N = 10 - 2 = 8; \; \alpha = 0.05$$

$$nn \frac{8(0.0183)(0.98167)}{(7)(0.15)^2(1.96)^{-2} + (0.0183)(0.98167)} = \frac{0.14372}{0.01822 + 0.01796} = 3.96 = 4$$

This means that a simple random sampling without replacement of four (4) public hospitals across the eight (8) Local Government Areas.

2.2 The Logistic Regression Model

The logistic regression model for multiple predictor variables is given as

$$ln\left(\frac{\pi(y)}{1-\pi(y)}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots \beta_k X_{ki} \qquad (2)$$

This looks like a classical multiple linear regression model; but in this case, the underlying distribution is binomial distribution. Hence, the parameters cannot be estimated the same way as multiple linear regression model. Rather, they are usually estimated using the maximum likelihood estimation method.

The likelihood function is used to estimate the probability of observing the data given the unknown parameters β_0, β_1,… β_k. The likelihood is the probability that the observed values of response variable may be predicted using the observed values of the independent variables. The likelihood varies from 0 to 1, just like any other probabilities… practically, it is easier to work with the logarithm of the likelihood function and this is known as the log-likelihood.

Log-likelihood as used for inference testing when comparing several models and its values range from zero (0) to infinity (∞).

In logistic regression, binary outcomes and predictors are observed for the purpose of drawing inferences about the probability of an event in the population. Suppose in a population in which sampling is done each individual has the sample probability π that an event occurs. For each individual in the sample of size n, $Y_i = 1$ indicates that an event occurs for the ith subject, otherwise $Y_i = 0$. The observed data are $Y_1, Y_2, \dots, Y_n$ and $X_1, X_2, \dots, X_n$.

The likelihood function is given by

$$L = \prod_{i=1}^{n} \{\pi(y)^{Y_i} [1 - \pi(y)]^{1-Y_i}\}$$

$$= \pi(y)^{\sum_{i=1}^{n} Y_i} [1 - \pi(y)]^{n-\sum_{i=1}^{n} Y_i} \quad (3)$$

$$\ln L = \sum_{i=1}^{n} Y_i \ \ln \pi(y) + (n - \sum_{i=1}^{n} Y_i) \ln [1 - \pi(y)] \quad (4)$$

which,

$$\frac{\pi(y)}{1 - \pi(y)} = e^{\beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki}} \quad (5)$$

$$\pi(y) = \frac{e^{\sum_{i=0}^{n} \beta_i X_i}}{1 + e^{\sum_{i=0}^{n} \beta_i X_i}} \quad (6)$$

In this case, X_i's are the socio-demographic characteristics of patients. β_i's are the parameters of the model. It should be noted that the parameters of the model are estimated analytically through directive procedure, since there is no closed form of the maximum likelihood function. Newton Raphson method may be used to estimate the parameters of the model.

2.3 Odds ratio

The Log-odd of the dependent variable is given as $\ln\left(\frac{\pi(y/x)}{1-\pi(y/X)}\right)$

The odds ratio (OR) is defined as the ratio of the probability of the event occurring to the probability of it not occurring, given as follows:

$$O\,R = \frac{\pi(y)}{1-\pi(y)} = e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k} \quad (7)$$

OR is interpreted that, when an independent variable X_i increased by one unit $(X_i + 1)$, with all other predictors kept constant, the odds of the dependent variable increase by a factor $e^{\beta_i X}$, which is called the odd ratio (OR) and ranges from zero (0) to infinity (∞)

2.4 Chi-Square Goodness-of-fit Tests

Instead of using the multiple coefficients of determination (R^2) as the statistic to measure overall fit of the model in logistic regression, the deviance between observed values from expected values is used. Let

$$\hat{y}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \beta_k X_{ki})}{1 + \exp \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \beta_k X_{ki}} \quad (8)$$

The standardized residuals can be expressed as

$$\tau_i = \frac{y_i - \hat{y}_i}{\sqrt{\hat{y}_i (1 - \hat{y}_i)}} \quad (9)$$

where y_i is the observed dependent variable for the ith subject and $\hat{y}_i$ is the predicted dependent variable.

The standard deviation of the residual is $\sqrt{\hat{y}_i(1 - \hat{y}_i)}$

The chi-square statistics can then be formed as

$$\chi^2 = \sum_{i=1}^{n} \tau_i^2 \quad (10)$$

This statistic follows a χ^2 – distribution with n – k – 1 degrees of freedom. The computed χ^2 is less than the tabulated χ^2, then the model fits well to the data.

2.5 Hosmer – Lemeshow test

This test examines whether the observed proportions of events are similar to the predicted probabilities of occurrence in subgroups of the model population. The test is performed by dividing the predicted probabilities into deciles (10 groups based on percentages ranks) and then computing a Pearson Chi-square that compares the predicted to the observed frequencies, (Park, 2013).

The list statistic is given as

$$H = \sum_{g=1}^{G}\left\{\frac{(O_{1g}-E_{1g})^2}{E_{1g}} + \frac{(O_{0g}-E_{0g})^2}{E_{0g}}\right\} \quad (11)$$

where O_{ig} and E_{ig} are the observed and expected events for the gth decile group. O_{0g} and E_{0g} are the observed and expected number of non-events. The test statistic approximately follows a chi-square distribution with G – 2 degrees of freedom (10 – 2 = 8).

The decision rule is that if p-value is greater than 0.05, accept H_0 that the model fits adequately to the data and if p-value < 0.05, H_0 is rejected.

2.6 Statistical Significance Test

1.Likelihood ratio test: The overall fit of the model with co-efficient can be tested using a likelihood ratio test which tests the null hypothesis.

$H_0: \beta_0 = \beta_1 = \cdots = \beta_k = 0$

against

$H_1: \beta_0 \neq \beta_1 \neq \ldots \neq \beta_k \neq 0$ for at least one

To test this, the deviance when only intercept is included in the model (-2 long – likelihood of the null model) is compared to the deviance when the p-variables are added in the model (-2 log-likelihood of the given model). The difference of these two yields the test statistic, ϱ, which is a chi-square statistic with k degrees of freedom, the (Bewick *et al.*, 2005). Thus,

$$\varrho = \chi^2 = (-2 loglikelihood\ of\ the\ null\ model) - (-2\ loglikelihood\ of\ given\ model)$$

$$\varrho = -2log\frac{likelihood\ of\ null\ model\ (L_0)}{likelihood\ of\ given\ model\ (L_1)} = -2log\frac{L_0}{L_1} \quad (12)$$

The decision is that if calculated χ^2 is greater than tabulated χ^2, reject H_0, otherwise accept.

2.6,1 Wald Statistic

The Wald statistic is used to test the significance of individual parameter estimates predictor variables) in the model; (Park, 2013). The Wald statistic is given by

$$W = \left[\frac{\widehat{\beta_i}}{SE(\widehat{\beta_i})}\right]^2 \quad (13)$$

$W \sim \chi^2$ with 1 degree of freedom.

W is used to test the hypothesis

$H_0: \beta_i = 0$ against

$H_1: \beta_i \neq 0$; $i = 0, 1, 2, \ldots, k$

If the p-value is less than 0.05, H_0 is rejected and it is concluded that the variable is statistically significant in the model.

2.7 Predictive Accuracy

The classification table is adopted to evaluate the predictive accuracy of the logistic regression model. In the whole, the observed values for the dependent outcome and the predicted values are classified. For example, if a cut-off value is 0.5, all predicted values above 0.5 can be classified as predicting an event, and all below 0.5 as not predicting the event. Then, a 2 x 2 table of data can be constructed with dichotomous observed outcomes and dichotomous predicted outcomes.

If the logistic regression model has a good fit, it is expected that many counts in n_{11} and n_{22} cells will be observed, and few in the n_{12} and n_{21} cells. Sensitivity is considered ad ${n_{11}}/{n_{11}+n_{12}}$ and specificity is ${n_{22}}/{n_{21}+n_{22}}$. Higher sensitivity and specificity indicate a better fit of the model.

Table 1: Sample classification table

Observed	Predicted

	1	0
1	n_{11}	n_{12}
0	n_{21}	n_{22}

3.0 Results and Discussion

3.1: Descriptive Statistics

Table 2 indicates that out of 1581 who tested HIV-negative , 1576 representing 99.7% had no family history, while 5 (0.3%) had a family history of HIV infection. Also, out of 1030 patients who were tested positive, 890 (86.4%) had no family history of HIV infection, whereas 140 (13.6%) reported a family history of HIV infection..

Table 3 shows out of a total of 1581 patients, who tested negative, 1036 representing 65.5% were females, while 545 representing 34.5 % were males. On the other hand, of 1030 patients that tested positive, 523 (50.8%) were females and 507 (49.2%) were males. So, there were more female infected patients than male within the period of study.

From Table 4, it is observed that a total of 1581 patients tested negative, while 1030 tested positive. From those whose status were negative, 647 (40.9%) were married, 775 (49.0%) were singles, 32 (2.0%) were divorced, 101 (6.4%) were widows, while 26 (1.6%) were separated from their spouses. Also, out of those who tested positive, 482 (46.8%) were married, 456 (44.3%) were singles, 20 (1.9%) were divorcees, 53 (5.1%) were widows and 19 (1.8%) were separatedMarried and single patients constituted the largest proportions of both HIV-negative and HIV-positive patients because they represented the largest groups in the study population. These descriptive frequencies alone do not indicate that marital status increases the likelihood of HIV infection.

Table 2: HIV infection status with family history

			FAMILY HISTORY		
			No	Yes	Total
HIV STATUS	-ve	Count	1576	5	1581
		% within HIV STATUS	99.7%	0.3%	100.0%
	+ve	Count	890	140	1030
		% within HIV STATUS	86.4%	13.6%	100.0%
Total		Count	2466	145	2611
		% within HIV STATUS	94.4%	5.6%	100.0%

Table 3: HIV infection status with sex of patients

			SEX		
			Female	Male	Total
HIV STATUS	-ve	Count	1036	545	1581
		% within HIV STATUS	65.5%	34.5%	100.0%
	+ve	Count	523	507	1030
		% within HIV STATUS	50.8%	49.2%	100.0%
Total		Count	1559	1052	2611
		% within HIV STATUS	59.7%	40.3%	100.0%

Table 4: HIV infection status and marital status

			MARITAL STATUS					
			Married	Single	Divorced	Widowed	Separated	Total
HIV STATUS	-ve	Count	647	775	32	101	26	1581
		% within HIV STATUS	40.9%	49.0%	2.0%	6.4%	1.6%	100.0%
	+ve	Count	482	456	20	53	19	1030
		% within HIV STATUS	46.8%	44.3%	1.9%	5.1%	1.8%	100.0%
Total		Count	1129	1231	52	154	45	2611
		% within HIV STATUS	43.2%	47.1%	2.0%	5.9%	1.7%	100.0%

From Table 5, patients who had secondary school experience had the highest percentage of testing negative (932 out of 1581). This is seconded by those were in tertiary level (343 out of 1581). Thereafter, it was followed by patients with primary level of education, who were 199 in number. For those who tested positive, patients with the same secondary level of education were the highest in number of 566 (55%), while this was followed by those in tertiary level with a total number of 284 (27.6%), before it was followed by those that had primary level of education with a total of 116 (11.3%) patients. Patients with secondary education constituted the largest proportion of both HIV-negative and HIV-positive participants because this educational category represented the majority of the study population. The descriptive statistics alone do not imply that higher educational attainment increases the risk of HIV infection

Table 5: HIV infection status with levels of education

			LEVEL OF EDUCATION				
			Informal	Primary	Secondary	Tertiary	Total
HIV STATUS	-ve	Count	107	199	932	343	1581
		% within HIV STATUS	6.8%	12.6%	59.0%	21.7%	100.0%
	+ve	Count	64	116	566	284	1030
		% within HIV STATUS	6.2%	11.3%	55.0%	27.6%	100.0%
Total		Count	171	315	1498	627	2611
		% within HIV STATUS	6.5%	12.1%	57.4%	24.0%	100.0%

From Table 6, 692 (43.8%) self-employed patients test negative and was the highest, while 378 patients who are unemployed also tests negative as the second highest. This is followed by those who work in private firms with a total number of 337 representing 21.3%. The least are the public servants with 174. Looking at positive infection status, Self-employed patients constituted the largest proportion of HIV-positive patientsas they record 407 infections, which is equivalent to 39.5% of the total patients infected. This has patients who work in private firms as second most affected set, with a total of 264 patients representing 25.6% of patients with positive status. The third most affected are the unemployed persons with a total of 245 (23.8%) patients. The least are the civil servants or public servants.

Table 6: HIV infection status with employment status

			EMPLOYMENT STATUS				
			Self-employed	Govt./Civil Servant	Private	Unemployed	Total
HIV STATUS	-ve	Count	692	174	337	378	1581
		% within HIV STATUS	43.8%	11.0%	21.3%	23.9%	100.0%
	+ve	Count	407	114	264	245	1030
		% within HIV STATUS	39.5%	11.1%	25.6%	23.8%	100.0%
Total		Count	1099	288	601	623	2611
		% within HIV STATUS	42.1%	11.0%	23.0%	23.9%	100.0%

From Table 7, it is observed that 853 representing 54% of patients who are HIV negative live in the urban area, while 728 representing 46% of patients who are HIV negative reside in the rural area. Alternatively, 592 representing 57.5% of the patients who are HIV positive live in urban area, whereas, 438 representing 42.5% of patients who are HIV positive live in rural area. Urban residents constituted a slightly higher proportion of both HIV-negative and HIV-positive patients than rural residents. These descriptive statistics alone do not establish an association between place of residence and HIV infection

Table 7: HIV infection status with residence

			RESIDENCE		
			Urban	Rural	Total
HIV STATUS	-ve	Count	853	728	1581
		% within HIV STATUS	54.0%	46.0%	100.0%
	+ve	Count	592	438	1030
		% within HIV STATUS	57.5%	42.5%	100.0%

Total	Count	1445	1166	2611
	% within HIV STATUS	55.3%	44.7%	100.0%

From Table 8, patients within the age range of 11 to 50 years were predominantly not infected with HIV/AIDS, while patients with the same age groups excluding 11-20 years were significantly infected with HIV/AIDS. A total of 791 representing 76.8% of patients within the age group of 21-50 years were mostly infected with HIV/AIDS.

3.2 *The Binary Logistic Regression Analysis*

After the descriptive statistics exploration of the association of sociodemographic characteristics with HIV infection status of patients attending public hospitals in Akwa Ibom North-West Senatorial district, the data were further analysed using binary logistic regression Table 9 shows the tests for statistical significance of predictors (socio-demographic factors) to the response variable (HIV infection status of patients).

Table 8: HIV infection status with age

			AGE_GROUP							Total
			1-10	11-20	21-30	31-40	41-50	51-60	>60	
HIV STATUS	-ve	Count	52	252	489	387	269	83	49	1581
		% within HIV STATUS	3.3%	15.9%	30.9%	24.5%	17.0%	5.2%	3.1%	100.0%
	+ve	Count	50	77	273	303	215	77	35	1030
		% within HIV STATUS	4.9%	7.5%	26.5%	29.4%	20.9%	7.5%	3.4%	100.0%
Total		Count	102	329	762	690	484	160	84	2611
		% within HIV STATUS	3.9%	12.6%	29.2%	26.4%	18.5%	6.1%	3.2%	100.0%

Table 9: Test for the overall logistic regression model

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.428	.040	114.515	1	.000	.651

Table 10: Goodness of fit test

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	10.661	8	.222

Table 9 indicates that the fitted logistic regression model is adequate and significant since the computed p-value, 0.000 is less than 0.05 significance level.

Table 10 shows the result of the Hosmer and Lemeshow test for goodness of fit of the model. The Hosmer–Lemeshow goodness-of-fit test was not statistically significant ($\chi^2 = 10.661$, df = 8, p = 0.222), indicating that there was no evidence of poor model fit. Therefore, the fitted logistic regression model adequately described the observed dataFrom Table 11, family history is significant in the model as the p-value is less than the significance level of 0.05. Patients with the family history of HIV infection have 93.31 times the odds of being HIV positive compared with those without family history, provided all other variables are controlled. This finding indicates that family history of HIV infection was independently associated with HIV-positive status after adjusting for other socio-demographic variablesThe Table also shows that Sex is statistically significant in the model since p-value of 0.000 is less than 0.05. Male patients have 1.774 times (77.4% higher) odds of being HIV positive than female patients. Also, being married (1) is significant at 10% level of significance (p-value, 0.079<0.10) while being divorced is significant at 5% level (0.047<0.05). Divorced patients had 32.9% lower odds of HIV infection compared with the reference category (separated patients)

Table 11: Test for significance of parameter

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Family History(1)	4.536	.496	83.538	1	.000	93.308	35.277	246.801
	Sex(1)	.573	.089	41.803	1	.000	1.774	1.491	2.110
	Marital Status			8.859	4	.065			
	Marital Status(1)	.192	.109	3.085	1	.079	1.211	.978	1.500
	Marital Status(2)	-.075	.315	.057	1	.811	.927	.500	1.719
	Marital Status(3)	-.399	.201	3.940	1	.047	.671	.452	.995
	Marital Status(4)	-.184	.347	.281	1	.596	.832	.421	1.642
	Level of Education			17.217	3	.001			
	Level of Education(1)	.208	.235	.783	1	.376	1.231	.777	1.951
	Level of Education(2)	.559	.202	7.658	1	.006	1.748	1.177	2.597
	Level of Education(3)	.737	.212	12.092	1	.001	2.089	1.379	3.165
	Employment Status			6.096	3	.107			
	Employment Status(1)	-.094	.147	.411	1	.522	.910	.683	1.214
	Employment Status(2)	.231	.110	4.410	1	.036	1.260	1.016	1.564
	Employment Status(3)	.047	.124	.146	1	.703	1.048	.822	1.337
	Residence(1)	-.201	.088	5.176	1	.023	.818	.688	.973
	Age_group			51.625	6	.000			
	Age_group(1)	.084	.390	.046	1	.830	1.087	.506	2.337
	Age_group(2)	.783	.376	4.323	1	.038	2.187	1.046	4.574
	Age_group(3)	1.142	.383	8.905	1	.003	3.134	1.480	6.637
	Age_group(4)	1.251	.390	10.268	1	.001	3.495	1.626	7.513
	Age_group(5)	1.531	.416	13.571	1	.000	4.625	2.048	10.447
	Age_group(6)	1.395	.439	10.091	1	.001	4.034	1.706	9.540
	Constant	-2.327	.429	29.454	1	.000	.098		

a. Variable(s) entered on step 1: FAMILY HISTORY, SEX, MARITAL STATUS, LEVEL OF EDUCATION, EMPLOYMENT STATUS, RESIDENCE, AGE_GROUP.

For the educational level, patients being in primary and secondary schools are statistically significant in the model, as their p-values, 0.006 and 0.001 are less than 0.05. Patients with primary education had 74.8% higher odds of HIV infection than those with tertiary education, while those with secondary education had approximately twice the odds of HIV infection. Patients with secondary level of education have 2.089 times (108.9%higher odds) the odds of being HIV positive than patients in tertiary level. In addition, patients who are government employed are statistically significant in the model. Those patients have 1.26 times (26% higher odds) the odds of being HIV positive than patients who are unemployed. Still from Table 11, urban residence is statistically significance in the model, since 0.023 < 0.05. Patients, who live in urban area have 0.818 times (18.2% less odds) the odds of being HIV positive than patients who live in rural area. Finally, all the age groups, except age group 1-10 years, are statistically significant, since all their p-values are less than 0.05. Patients with age group 11-20, 21-30, 31-40, 41-50 and 51-60 years have 2.187, 3.134, 3.495, 4.625 and 4.034 times the odds of being HIV positive than patients with age group of greater than 60 years.

Table 12 shows that the binary logistic regression model correctly classifies 91.3% of HIV-negative patients and 34.4% of HIV-positive patients, with an overall classification accuracy of 69%. "Although the model achieved an overall classification accuracy of 68.9%, it demonstrated substantially higher specificity (91.3%) than sensitivity (34.4%), indicating that it was considerably better at identifying HIV-negative patients than HIV-positive patients.

Table 12: Classification

			Predicted HIV STATUS		
	Observed		-ve	+ve	Percentage Correct
Step 1	HIV STATUS	-ve	1444	137	91.3
		+ve	676	354	34.4
	Overall Percentage				68.9

a. The cut value is .500

3.3: Discussion of Results

From the descriptive statistics of the data, it has been observed that patients with family history of HIV infection have higher frequency of patients being HIV-positive. The logistic regression analysis shown in Table 12 confirms this too (AOR = 93.31, $p<0.05$). This agrees with the study conducted by Kefale *et al.* (2023) in Sodo town of Southern Ethiopia using clinical attendants as subjects. The work found that the risk of HIV infection among family members of index cases was highThe observed association may reflect shared environmental, behavioural, or familial factors rather than direct transmission through routine family contact.. Also, results show that male patients (1.774, $p<0.05$) have higher odds of being infected with HIV than females, while the descriptive statistics show that more females are HIV-positive than male patients, with only a difference of 16 females. This result is in line with the works of Patel *et al.* (2022), which recorded high HIV infection among both male and females in Sub-Saharan

Africa. A retrospective study of HIV hospitalized patients in Nigeria revealed that male sex was associated with significantly worse HIV outcomes, (Agada, *et al*., 2011). Although this result is in deviance with many population-based studies in sub-Saharan Africa and Nigeria in particular, which have reported higher prevalence of HIV among females, the observed difference may be attributed to the hospital-based nature of the present study, where health care utilization patterns delayed health-seeking behaviour among males, referral practices, and the characteristics of patients attending the selected hospitals differ from those of the general population. For marriage status, it is discovered that divorced patients (0.671, $p<0.05$) tend to be more likely to be infected with HIV than separated persons. This is in agreement with the works of Tenkorang (2014) that divorced women had higher risks of HIV infection compared to never-married women. This may be due to the fact that marriage dissolutions can expose partners to economic and social instability and more involvements in multiple sex partners. Moreover, the study confirms that patients with both primary and secondary levels of education (AOR=1.75, 2.09, $p<0.05$ respectively) have significantly higher odds of being HIV-positive. The observed association may reflect differences in healthcare-seeking behaviour, HIV testing uptake, occupational exposure, or other unmeasured socio-economic factors rather than education itself increasing susceptibility to HIV infection. It may also stem from the fact that those at that level may not have the awareness. This result is corroborated by the study undertaken by Musa, *et al*. (2026). It is also observed from the analysis that government employed patients (AOR = 1.026, $p<0.05$) have significantly higher odds of being HIV-positive. This may be due to awareness on the part of the workers to visit hospitals for HIV tests; it may also be due to the government employed people being exposed to occupational mobility, increased social interaction, or other behavioural and socio-economic factors. This is in agreement with the study conducted by Nastiti *et al.* (2024) in Indonesia, which found that employment status is one of the risk factors of HIV infection. Also, the analysis still indicates that patients living in urban areas (AOR = 0.818, $p<0.05$) have lesser odds of being HIV-positive. The lower odds observed among urban residents may reflect better access to HIV prevention programmes, testing services, and healthcare facilities, although additional studies are required to confirm this explanationThis is in line with the works of Olubayo *et al.* (2025), which indicated that adults living in rural areas have a higher odd of HIV infection than those living in urban areas. Another factor is age group. It is found that patients whose age groups are 11-20, 21-30, 31-40, 41-50 and 51-60 years (2.187, 3.134, 3.495, 4.625 and 4.034; $p<0.05$) have higher odds of being HIV-positive.

4.0 Conclusion

This study investigated the influence of selected socio-demographic characteristics on the HIV infection status of patients attending selected public hospitals in Akwa Ibom North-West Senatorial District using multiple binary logistic regression analysis. The findings revealed that family history of HIV infection, sex, marital status, educational level, employment status, place of residence, and age were significant predictors of HIV infection status. Specifically, patients with a family history of HIV infection, male patients, those with primary or secondary education, government-employed individuals, and patients aged 11–60 years had significantly higher odds of testing HIV-positive, whereas patients residing in urban areas had significantly lower odds of HIV infection than their rural counterparts. Although divorced patients differed significantly from the reference marital status category, the interpretation should be made with respect to the selected reference group. The study concludes that socio-demographic characteristics are important determinants of HIV infection status among patients attending public hospitals in Akwa Ibom North-West Senatorial District. These findings highlight the need for targeted HIV prevention, screening, and health education programmes that prioritize high-risk population groups identified in this study. Furthermore, strengthening community-based awareness campaigns, expanding access to HIV testing and counselling services, particularly in rural communities, and implementing evidence-based public health interventions tailored to vulnerable socio-demographic groups will contribute to reducing the burden of HIV infection in the study area.

5.0 References

Abogoh, L. O., Thomas, R. E., Peter, E. S., Asuquo, A. E., & Asuquo, C. A. (2025). The double burden of HIV/AIDS on livelihoods in Akwa Ibom State, Nigeria. *International Journal of Finance, Management and Governance*, 1, 3, pp. 36–43.

Agaba, P. A., Digin, E., Makai, R., Apena, L., Agbaji, O. O., Idoko, J. A., Murphy, R., & Kanki, P. (2011). Clinical characteristics and predictors of mortality in hospitalized HIV-infected Nigerians. *Journal of Infection in Developing Countries*, 5, 5, pp. 377–382.

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care*, 9, 1, pp. 112–118. http://dx.doi.org/10.1186/cc3045

Gaughan, C., Gaboyan, S., Tovar, C., Prado, M., Cappiello, M., & Burgos, J. L. (2026). Increased prevalence of people living with HIV among unhoused patients in Tijuana. *International Journal of STD & AIDS*, 37, pp. 658–661.

Gorfu, G. M., Sime, S. T., & Beyene, J. A. (2026). Prevalence, associated risk factors, knowledge, attitude, and practice of HIV/AIDS among patients visiting Metema Yohanness Primary Hospital, Northwest Ethiopia. *Journal of the International Association of Providers of AIDS Care*, 25.

Howard, N. (2024). Statistical approach to enhancing performance of logistic regression model: Application to HIV/AIDS data. *International Journal of Research and Innovation in Applied Science*, VIII, XII, pp. 230–245.

Iseh, M. J., Usoro, A. J., Ekong, N. P., & Ukpe, I. (2022). Juxtaposing vertically transmitted infections (VTIs) and the spread of HIV/AIDS in a typically infection prevalent region in Nigeria. *Journal of the Nigerian Society of Physical Sciences*, 4, 1, pp. 99–104.

Kefale, A., Daka, K., Abebe, A., Haile, D., Paulos, K., Sherfa, A., Addis, A., Gunta, M., Ayza, A., & Wolde, J. (2023). Prevalence of HIV seropositive status and associated factors among family members of index cases of antiretroviral clinical attendants in Sodo Town, Southern Ethiopia. *PLoS ONE*, 18, 2, e0280571. https://doi.org/10.1371/journal.pone.0280571

Lawan, U. M., Envuladu, E. A., & Abubakar, S. (2016). Does awareness of status and risks of human immunodeficiency virus impact risky transmission behavior among infected adolescents? A case study of clients attending an antiretroviral therapy (ART) clinic in Kano, Kano State, Nigeria. *Indian Journal of Community Medicine*, 41, pp. 126–132.

Mendez, A. D., Neelamegam, M., & Griner, S. B. (2023). Health care provider discussions regarding HIV/sexually transmitted infection risk factors and associations with HIV/sexually transmitted infection screening among men. *Archives of Sexual Behavior*, 52, pp. 2111–2121.

Musa, D. A., Okereke, T. G., Aremu, K. H., Umar, Y., Sule, E. B., Majiya, H., Aigboeghian, E. A., & Morka, M. C. (2026). Socio-demographic determinants of HIV prevalence across Nigerian regions. *Discover Public Health*, 23, 598. https://doi.org/10.1186/s12982-026-01917-y

Nastiti, A. A., Trihariri, M., Pradanie, R., Nursalam, N., Qur'raniati, N., Hutton, N., Syulthoni, Z. B., & Arfin, H. (2024). Socio-demographic factors and their association with HIV risk behaviours among Indonesian females aged 15–24 years: A nationwide study. *Journal of Pediatric Nursing*, 79, pp. e170–e176.

National Agency for the Control of AIDS(NACA). (2024). *HIV prevalence data by states in Nigeria: Spectrum estimates*. Reported in *Punch Newspapers*, *Premium Times*, and *Tribune Online* (April 2025).

Olubayo, L. A. I., Mathema, T., Kabudula, C. K., Micklesfield, L. K., Mohamed, S. F., Kisiangani, I., Ntimane, C. B., Choma, S. S., Houle, B., Hazelhurst, S., Crowther, N., Tollman, S., Tluway, F. D., Ramsay, M., & Gomez-Olive, F. X. (2025). The prevalence, incidence, and sociodemographic risk factors of HIV among older adults in sub-Saharan Africa (AWI-Gen): A multicentre, longitudinal cohort study. *Lancet Healthy Longevity*, 6, 100690. https://doi.org/10.1016/j.lanhl.2025.100690

Park, H. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Korean Journal of Academy of Nursing*, 43, 2, pp. 154–164.

Patel, C. J., Claypool, K. T., Chow, E., Chung, M., Mai, D., Jessie, C., & Bendavid, E. (2022). The demographic and socio-economic correlates of behaviour and HIV infection status across sub-Saharan Africa. *Communications Medicine*, 2, 104. https://doi.org/10.1038/s43856-022-00170-z

Tekle, G. (2019). Assessment of socio-economic, demographic and health factors that influences the survival/death status of HIV positive people under ART follow-up at Wolaita Sodo Referral Hospital (Ottona), Ethiopia. *Journal of Tropical Diseases & Public Health*, 7, pp. 1–8.

Tenkorang, E. Y. (2014). Marriage, widowed, divorce and HIV risks among women in sub-Saharan Africa. *International Health*, 6, 1, pp. 46–53.

Tugga, H. A., Ogunmola, A. O., Bamigbala, O. A., & Ahmad, S. S. (2026). Estimation of binary logistic regression using three links function (logit, probit, and complementary log-log) in assessing the factors that influence HIV. *Asian Journal of Science, Technology, Engineering, and Art*, 4, 3, pp. 373–376.

Utip, U. (2025, March 3). HIV prevalence: Akwa Ibom ranks highest with 5.5%. *Tribune Newspapers*. https://tribuneonlineng.com/hiv-prevalence-akwa-ibom-ranks-highest-with-5-5-percent/

Declarations

Ethical Approval and Consent to Participate

Ethical approval for this study was obtained from the appropriate Health Research Ethics Committee before data collection commenced. Permission to access hospital records was obtained from the management of the participating hospitals. The study was conducted in accordance with the principles of the Declaration of Helsinki. Patient information was anonymized to ensure confidentiality, and no personally identifiable information was collected or reported.

Consent for Publication

Not applicable.

Availability of Data and Materials

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Data are not publicly available because they contain information obtained from hospital records and are subject to institutional confidentiality requirements.

Competing Interests

The authors declare that they have no competing interests.

Funding

The authors received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors for the conduct of this study.

Authors' Contributions

Unyime Patrick Udoudo conceived and designed the study, collected and analysed the data, interpreted the results, and drafted the manuscript. Eduma E. Essien contributed to the study design, literature review, interpretation of the findings, and critical revision of the manuscript. Emmanuel J. Ekpenyong supervised the statistical analysis, validated the results, reviewed and edited the manuscript, and approved the final version for publication. All authors read and approved the final manuscript.

Acknowledgements

The authors sincerely appreciate the management and staff of the participating public hospitals in Akwa Ibom North-West Senatorial District for granting access to the hospital records used in this study. The authors also acknowledge all individuals whose support contributed to the successful completion of this research.

Authors' Information

Not applicable.

Consent to Participate

As this study involved retrospective analysis of anonymized hospital records, informed consent from individual patients was waived by the approving ethics committee in accordance with applicable ethical guidelines.

Data Confidentiality

All patient records were anonymized before analysis. Confidentiality and privacy were maintained throughout the study, and all data were used solely for research purposes.

Publisher's Note

The authors are solely responsible for the content of this manuscript. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.