FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 7, July, 2025, pp 240 – 244
DOI: https://doi.org/10.33003/fjs-2025-0907-3541

A QUANTILE REGRESSION ANALYSIS OF THE RELATIONSHIP BETWEEN DEMOGRAPHIC FACTORS AND AGE DISTRIBUTION AMONG HIV/AIDS PATIENTS IN JOS, PLATEAU STATE

***[1]Alade, Segun Peter, [1]Uwah Anefiok, [1]Su'an Luka and [2]Adaobi Ihionu**

[1]Department of Statistics, Plateau State Polytechnic, Barkin-Ladi, Nigeria
[2]Department of Statistics, University of Jos, Nigeria

*Corresponding authors' email: segunpeteralade@yahoo.com

ABSTRACT

Human Immunodeficiency Virus (HIV) remains a major public health concern in sub-Saharan Africa, with demographic factors influencing both vulnerability and disease progression. This study employs a quantile regression framework to explore how age interacts with gender, educational attainment, marital status, and employment status among HIV-positive patients in Jos, Plateau State, Nigeria. We conducted a retrospective cross-sectional analysis of 667 available patient records from local treatment centres, fitting quantile regression models at the 25th, 50th, 75th, and 90th percentiles ($\tau = 0.25, 0.50, 0.75, 0.90$) using R (v4.3.2). Descriptive statistics indicated a right-skewed age distribution (mean = 33.0 years; skewness = 2.28). Normality tests (Kolmogorov–Smirnov, Shapiro–Wilk) confirmed significant departures from normality across most subgroups, justifying the quantile approach. Results revealed that female gender is significantly associated with younger age at diagnosis at the 25th–75th percentiles ($p < 0.01$), whereas lower educational status becomes a significant risk factor at the 50th–90th percentiles ($p < 0.001$). In contrast, marital status and employment status exhibited no significant associations with age across any quantiles. These findings show the disproportionate vulnerability of young women (approximately ages 26–40) and less-educated adults to HIV infection. Recommendations include gender-targeted prevention strategies, particularly a comprehensive sexual education and empowerment initiatives for young women and broader application of quantile regression methods in epidemiological research to capture distributional heterogeneity.

Keywords: HIV/AIDS, Quantile Regression, Age Distribution, Demographic Factors, Vulnerability, Disease Progression

INTRODUCTION

Human Immunodeficiency Virus (HIV) and Acquired Immunodeficiency Syndrome (AIDS) continue to pose significant public health challenges globally, particularly in sub-Saharan Africa, which accounts for nearly two-thirds of the global HIV burden (UNAIDS, 2023). Although advancements in treatment have improved the quality of life and life expectancy of people living with HIV, the disease remains incurable and is often associated with persistent stigma and discrimination (WHO, 2022). Unlike the recent global outbreaks of Ebola or COVID-19 that drew widespread attention, HIV/AIDS represents a long-standing epidemic that has silently reshaped health and demographic profiles over the decades.

HIV/AIDS does not discriminate by age, sex, race, or socioeconomic status, but research has shown that certain demographic variables significantly influence both vulnerability to infection and disease progression (CDC, 2023). Age, in particular, has emerged as a critical factor in understanding HIV transmission patterns. Several studies have consistently demonstrated that intergenerational sexual relationships, especially those between adolescent girls and significantly older men, substantially contribute to the elevated HIV incidence among young women (Dellar, Dlamini & Karim, 2022; Birdthistle *et al.,* 2021; Beckford, 2018; Gregson *et al*., 2002; Harling *et al*., 2014). This form of "age-mixing" exacerbates the risk profile for young people, especially in resource-constrained settings with limited access to education and reproductive health services.

Demographic characteristics such as gender, educational attainment, marital status, and employment status interact with age to influence HIV risk. For instance, young women are more biologically susceptible to HIV infection and often face gender-based power imbalances that limit their ability to negotiate safe sex (Jewkes *et al*., 2019). Education has also been shown to serve as a protective factor. Recent meta-analyses suggest that higher educational attainment is consistently associated with reduced HIV prevalence, especially when complemented by comprehensive sexual health education (Chimbindi *et al*., 2021). Conversely, low educational levels are linked with reduced HIV knowledge and limited access to preventive resources (Pettifor *et al.,* 2018).

Similarly, employment status and economic security play critical roles in shaping health-seeking behaviour and access to treatment. However, recent findings indicate that while employment itself does not directly correlate with HIV prevalence, the nature and stability of employment can mediate HIV-related outcomes by influencing social mobility and healthcare accessibility (Mabaso *et al*., 2021). Marital status is another complex factor. Although marriage is often perceived as a protective institution, studies show that HIV risk remains high among both married and unmarried individuals, depending on partner fidelity and access to health services (Gomez-Olive *et al*., 2020).

The link between age and HIV vulnerability is not merely biological but also behavioural and social. Age influences decision-making capacity, exposure to risk, and likelihood of engaging in preventive behaviour. Younger individuals are often less likely to use condoms consistently and may lack access to accurate HIV information, increasing their risk of infection (UNICEF, 2022). Therefore, understanding the interplay between age and other demographic variables is essential for designing targeted interventions that are age-sensitive and context-specific.

This study employs quantile regression analysis to examine how age interacts with selected demographic variables, gender, educational status, marital status, and employment status among HIV/AIDS patients in Jos, Plateau State, Nigeria. Unlike traditional regression techniques that focus on the mean of the response variable, quantile regression offers a deeper understanding of how these factors operate across different percentiles of the age distribution. This approach allows us to identify vulnerable subpopulations and inform more effective public health strategies.

MATERIALS AND METHODS

This study utilised a retrospective cross-sectional design, analysing data from HIV-positive individuals who accessed treatment in selected health facilities within Jos Metropolis, Plateau State, Nigeria, who are involved in the treatment of HIV/AIDS patients. The area is a major urban hub in North Central Nigeria and represents a significant region in the national HIV response due to its high population density and diverse demographics (NACA, 2022). Data were extracted from the medical records of 667 HIV-positive patients who were available using a structured data extraction form. Inclusion criteria required complete documentation of patient age, gender, educational status, marital status, and employment status. Age serves as the dependent variable while the independent variables are gender (male, female), educational status (non-formal, primary, secondary, tertiary), marital status (single, married, divorced, cohabiting, widowed) and employment status (student, employed, unemployed).

Categorical variables were coded using dummy variables for regression analysis. Given the presence of non-normality and outliers, which were confirmed by Kolmogorov-Smirnov and Shapiro-Wilk tests, Quantile Regression (QR) was employed rather than Ordinary Least Squares (OLS) for the analytical framework. QR allows the modelling of the conditional quantiles (e.g., median, quartiles) of the response variable and is particularly robust in the presence of heteroscedasticity and skewed distributions (Koenker, 2020; Huang *et al*., 2017). For a continuous response variable Y (in this case, age), and a vector of covariates X, the τ-th quantile regression model estimates the conditional quantile function using this formula:

$$Q_Y(\tau|X) = X^{\mathrm{T}}\beta_\tau \qquad (1)$$

Where:

$Q_Y(\tau|X)$ is the conditional quantile of Y given X,

β_τ is a vector of coefficients specific to the quantile $\tau \in (0, 1)$,

$X^{\mathrm{T}}\beta_\tau$ gives the linear combination of the predictors at quantile τ.

To estimate β_τ, the following minimisation problem is solved:

$$\hat{\beta}_\tau = arg \min_{\beta \in R^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i{}^{\mathrm{T}}\beta) \qquad (2)$$

Where: $\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$ is the check function,

$I(\cdot)$ is an indicator function,

$u = y_i - x_i{}^{\mathrm{T}}\beta$ is the residual for observation i.

The check function $\rho_\tau(u)$ assigns asymmetric weights to positive and negative residuals, enabling the estimation of various quantiles:

$$\rho_\tau(u) = \begin{cases} \tau u & u \geq 0 \\ (\tau - 1)u & u < 0 \end{cases} \qquad (3)$$

The advantage of this approach is that it provides insights into how the predictors affect not just the average age, but also different segments (quantiles) of the age distribution. For example, it allows the investigation of whether employment or educational status affects younger versus older HIV patients differently.

Quantile regression was performed at τ = 0.25, 0.50, 0.75, and 0.90 using R software version 4.3.2 and the *quantreg* package.

RESULTS AND DISCUSSION

Table 1 provides essential characteristics of the age distribution among the HIV/AIDS patients in our study sample. With a total of 667 participants, the average age was approximately 33 years (32.9955), with a standard deviation of 11.47 years, indicating considerable variability in the ages of individuals affected by HIV/AIDS. This variability is further highlighted by the coefficient of variation of 34.76%, which indicates that age values are widely spread out and suggests a moderately heterogeneous age distribution.

Table 1: Age Distribution Statistics

Item	Statistic
Count	667
Average	32.9955
Standard deviation	11.4689
Coeff. of variation	34.7589%
Minimum	0.0
Maximum	71.0
Range	71.0
Median	32.0
MAD	10.3782
Standard skewness	2.27747
Standard kurtosis	1.52544

The age range spans from newborns to elderly individuals (maximum age of 71), demonstrating that HIV/AIDS affects people across all life stages in the study population. The median age of 32 years, very close to the mean, suggests a relatively symmetric central tendency. However, when examining measures of distribution shape, the standardized skewness value of 2.28 falls outside the expected range of -2 to +2 for normal distributions, indicating a right-skewed age distribution with more younger patients and a tail extending toward older ages. This non-normal distribution justifies our analytical approach using quantile regression, which does not assume the normality of the response variable (Koenker, 2020; Huang *et al*., 2017).

Table 2: Test of Normality

Demographic characteristics	Variables	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	Df	Sig.
Employment Status	Student	0.099	57	0.200*	0.979	57	0.415
	Unemployed	0.085	511	0.000	0.985	511	0.000
	Employed	0.111	98	0.004	0.948	98	0.001
Educational Status	None	0.058	160	0.200*	0.993	160	0.597
	Primary	0.069	245	0.007	0.978	245	0.001
	Secondary	0.090	241	0.000	0.977	241	0.001
	Tertiary	0.135	20	0.200*	0.944	20	0.287
Gender	Male	0.115	196	0.000	0.963	196	0.000
	Female	0.117	470	0.000	0.970	470	0.000
Marital Status	Single	0.095	107	0.019	0.968	107	0.011
	Married	0.074	513	0.000	0.988	513	0.000
	Divorced	0.126	15	0.200*	0.929	15	0.260
	Co-habiting	0.260	2				
	Widowed	0.119	29	0.200*	0.975	29	0.704

Table 2 provides the validation for our methodological choice of quantile regression. Both Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted to assess the normality of age distributions across the selected demographic characteristics.

The results reveal that age distributions across most demographic subgroups significantly depart from normality. Specifically, only a few subgroups demonstrated normal age distributions, which include students, individuals with no formal education, those with tertiary education, and divorced individuals. All other subgroups, including the unemployed, employed, those with primary or secondary education, both males and females and most marital status categories, showed significant deviations from normality ($p < 0.05$).

These findings strongly support the decision to employ quantile regression analysis rather than ordinary least squares regression, as the latter assumes normally distributed residuals (Staffa *et al*., 2019). The presence of non-normality across most demographic subgroups, along with the potential presence of outliers and varying variances among factors, makes quantile regression particularly appropriate for capturing the complex relationships between demographic variables and age in this HIV/AIDS patient population, as emphasized by Huang *et al*. (2017) who noted that quantile regression offers more robust estimates when normality assumptions are violated.

Table 3: Coefficients and P-value for tau=0.25, 0.50, 0.75 and 0.90

	tau=0.25		*tau=0.50*		*tau=0.75*		*tau=0.90*	
	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value
Intercept	38.5000	*0.0000*	53.6667	*0.0000*	63.6667	*0.0000*	66.000	*0.0000*
Gender	-4.50000	*0.0006*	-6.0000	*0.0000*	-6.0000	*0.0002*	-2.5000	*0.1371*
Educational Status	-1.00000	*0.1041*	-3.0000	*0.0000*	-5.0000	*0.0000*	-4.5000	*0.0000*
Marital Status	0.50000	*0.3492*	-0.33333	*0.4943*	-0.3333	*0.7134*	-1.0000	*0.1652*
Employment Status	-1.50000	*0.1315*	-2.00000	*0.0131*	-1.0000	*0.5386*	-1.0000	*0.4843*

Table 3 presents the quantile regression coefficients and corresponding p-values at the 25th, 50th, 75th, and 90th percentiles, revealing the in-depth relationship between demographic factors and age among HIV/AIDS patients.

Gender emerges as a significant predictor of age of contact of HIV/AIDS at the 25th, 50th, and 75th percentiles, but not at the 90th percentile. The consistently negative coefficients (-4.5, -6.0, -6.0, and -2.5, respectively) indicate that females (coded as the lower level of gender) are significantly younger when contracting HIV/AIDS compared to males, particularly in the age range of 26 to 40 years (corresponding to the 25th through 75th percentiles). The finding implies that gender has a stronger association with age at the lower quantile than the upper. This finding aligns with UNAIDS (2017) data showing higher HIV prevalence among women in sub-Saharan Africa, where 61% of adults living with HIV/AIDS are female. The gender disparity observed in our study is also consistent with Dellar *et al*. (2022), who identified that adolescent girls and young women face disproportionate HIV risk due to biological vulnerability and gender-based power imbalances. The non-significance at the 90th percentile suggests that gender differences diminish among older patients (approximately 55 years and above).

Educational status shows an age-dependent pattern of significance. At the 25th percentile (younger patients), educational status has no significant impact ($p = 0.1041$). However, it becomes highly significant at the 50th, 75th, and 90th percentiles ($p < 0.0001$ for all), with increasingly negative coefficients (-3.0, -5.0, and -4.5, respectively). It indicates that educational status does not strongly influence the age of HIV/AIDS patients in the younger group (25th percentile), while at median and older age group (50th to 90th percentiles), education becomes a strong predictor – suggesting that has age increases, educational differences become more pronounced in influencing age distribution among patients. This supports existing literature showing that education increases self-efficacy and reduces HIV vulnerability, particularly in middle and older age groups (Lee *et al*., 2020; Chimbindi *et al*., 2021).

Notably, neither marital status nor employment status demonstrated significant associations with age across any percentile examined. The p-values for marital status (0.3492, 0.4943, 0.7134, and 0.1652) and employment status (0.1315, 0.0131, 0.5386, and 0.4843) were predominantly above the significance threshold of 0.05 (with one borderline exception for employment status at the 50th percentile). These findings indicate that individuals of all ages are equally vulnerable to HIV infection regardless of their marital or employment circumstances, contradicting potential assumptions about age-specific risks associated with these demographic factors. This complex relationship between marital status and HIV risk has been observed in previous studies from Southern Africa (Shisana *et al*., 2004), which found that HIV prevalence among married individuals did not differ significantly from unmarried individuals when controlling for other demographic factors. Similarly, Kordovski *et al*. (2017) found no significant interaction between age and employment status among HIV-infected individuals, supporting our findings regarding employment status.

CONCLUSION

This study applied quantile regression analysis to examine the influence of demographic variables on age distribution among 667 HIV/AIDS patients in Jos, Plateau State. Our findings reveal a complex interplay between demographic factors and age vulnerability to HIV infection. Gender emerged as a critical determinant, with females significantly more likely to contract HIV at younger ages (approximately 26-40 years) compared to their male counterparts, corroborating established epidemiological patterns observed throughout sub-Saharan Africa (UNAIDS, 2017). This gender disparity likely reflects both biological vulnerabilities and socio-structural inequalities that disproportionately expose young women to HIV risk, including limited economic opportunities, gender-based violence, and reduced negotiating power in sexual relationships (Jewkes *et al*., 2019).

Educational status demonstrated an age-dependent relationship with HIV vulnerability. While education did not significantly influence HIV risk among younger populations, its protective effect became increasingly pronounced among older age groups. This suggests that educational attainment functions as a cumulative protective factor that gains significance throughout adulthood, likely through mechanisms of enhanced health literacy, increased self-efficacy, and improved socioeconomic status (Chimbindi *et al*., 2021).

Contrary to some prevailing assumptions, neither marital status nor employment status showed significant associations with age-specific HIV risk patterns. This finding challenges simplistic narratives about HIV vulnerability and underscores the universal susceptibility across various social categories. The lack of age-specific effects for these variables suggests that HIV prevention strategies should not concentrate exclusively on particular marital or employment groups but should instead adopt inclusive approaches that recognize the complex, multifaceted nature of HIV risk factors.

RECOMMENDATIONS

Future research should explore additional demographic variables not included in this study, such as income level, ethnicity, religion, and geographic location (urban/rural). Incorporating these factors would provide a more comprehensive understanding of HIV vulnerability patterns and enable even more targeted prevention strategies. Additionally, longitudinal studies would help establish causal relationships between demographic factors and HIV risk over time.

Given the heightened vulnerability of young women to HIV infection, prevention efforts should prioritize gender-sensitive approaches that address both biological susceptibility and socio-structural factors that increase women's risk. These programs should include comprehensive sexual education, access to female-controlled prevention methods (such as female condoms and pre-exposure prophylaxis), and initiatives that empower women economically and socially. Particular attention should be directed toward females aged 15-40 years, where our study identified the most pronounced gender disparities.

The heterogeneous relationships revealed through quantile regression highlight the value of this methodology in HIV research. We recommend a wider application of quantile regression techniques in epidemiological studies to uncover complex patterns that might be obscured by traditional mean-based approaches.

REFERENCES

Birdthistle, I., Tanton, C., Tomita, A., de Graaf, K., Schaffnit, S. B., Tanser, F., ... & Hargreaves, J. R. (2021). Age-disparate partnerships and HIV incidence among adolescent girls and young women in rural South Africa: an HPTN 068 analysis. *Journal of the International AIDS Society, 24*(e25792). https://doi.org/10.1002/jia2.25792

Beckford, C. (2018). Age-disparate sexual relationships and HIV infection in Sub-Saharan Africa: A systematic review. *African Journal of AIDS Research, 17*(1), 1–9. https://doi.org/10.2989/16085906.2018.1432040

Chimbindi, N., Mthiyane, N., Birdthistle, I., Floyd, S., McGrath, N., Pillay, D., ... & Seeley, J. (2021). Persistently high HIV incidence among young women in rural KwaZulu-Natal, South Africa: A community-based cohort study. *The Lancet HIV, 8*(7), e429–e437.

Centre for Disease Control and Prevention. (2023). *HIV surveillance report, 2021*. U.S. Department of Health and Human Services. https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html

Dellar, R., Dlamini, N., & Karim, Q. A. (2022). Adolescent girls and young women: Key populations for HIV epidemic control. *Current Opinion in HIV and AIDS, 17*(3), 161–168.

Gomez-Olive, F. X., Houle, B., Angotti, N., Schatz, E., Reniers, G., Clark, S., ... & Tollman, S. (2020). HIV prevalence and incidence among adults in a rural South African community: The HAALSI cohort study. *Journal of the International AIDS Society, 23*(S1), e25586.

Gregson, S., Nyamukapa, C. A., Garnett, G. P., Wambe, M., Lewis, J. J., Mason, P. R., ... & Anderson, R. M. (2002). Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *The Lancet, 359*(9321), 1896–1903. https://doi.org/10.1016/S0140-6736(02)08780-9

Harling, G., Newell, M.-L., Tanser, F., Kawachi, I., Subramanian, S. V., & Bärnighausen, T. (2014). Do age-disparate relationships drive HIV incidence in young women? Evidence from a population cohort in rural KwaZulu-Natal, South Africa. *Journal of Acquired Immune Deficiency Syndromes, 66*(4), 443–451. https://doi.org/10.1097/QAI.0000000000000198

Huang, Q., Zhang, H., Chen, J., & He, M. (2017). Quantile regression models and their applications: A review. *Journal of Biometrics & Biostatistics, 8*(3), 1–6.

Jewkes, R., Jama-Shai, N., & Sikweyiya, Y. (2019). Ending the cycle of violence: Applying a gender lens to HIV prevention and care. *The Lancet, 393*(10189), 502–504.

Koenker, R. (2020). *Quantile regression* (Econometric Society Monograph Series). Cambridge University Press.

Kordovski, V. M., Woods, S. P., Verduzco, M., & Beltran, J. (2017). The effect of ageing and HIV disease on employment status and functioning. *Rehabilitation Psychology*, 64(4), 591-599.

Lee, N., Stucklin, S. B., Rodriguez, P. L., Faris, M. E., & Mukaka, I. (2020). Financial education for HIV-vulnerable youth, orphans, and vulnerable children: A systematic review of outcome evidence. *International Initiative for Impact Evaluation*.

Mabaso, M., Mlangeni, L., Jooste, S., Odemwingie, U., & Zuma, K. (2021). Socioeconomic and behavioural determinants of HIV infection among adults in South Africa. *International Journal of Environmental Research and Public Health, 18*(3), 1222.

National Agency for the Control of AIDS. (2022). *National HIV/AIDS Indicator and Impact Survey (NAIIS)*. Abuja: NACA.

Pettifor, A., Stoner, M., Pike, C., & Bekker, L.-G. (2018). Adolescent lives matter: Preventing HIV in adolescents. *Current Opinion in HIV and AIDS, 13*(3), 265–273. https://doi.org/10.1097/COH.0000000000000453

Shisana, O., Zungu, N. P., Toefy, Y., Simbayi, L., Malik, S., & Zuma, K. K. (2004). Marital status and HIV infection in South Africa. *South African Medical Journal*, 94(7), 537-543.

Staffa, S. J., Kohane, D. S., & Zurakowski, D. (2019). Quantile regression and its applications: A primer for anesthesiologists. *Anesthesia & Analgesia, 128*(4), 820–830.

UNAIDS (2023). *Global AIDS update 2023 – The path that ends AIDS*. Joint United Nations Programme on HIV/AIDS. https://www.unaids.org

UNICEF. (2022). *Young people and HIV – The missing link*. United Nations Children's Fund.

World Health Organization. (2022). *Global health sector strategy on HIV, 2022–2030*. https://www.who.int/publications/i/item/9789240053779