Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 3(3) May 2025
DOI: https://dx.doi.org/10.4314/jobasr.v3i3.20

Survival Analysis of Data on Human Immunodeficiency Virus (HIV) Infection among Children Aged One to Ten Years at General Hospital Minna, Niger State, Nigeria

Oguche, S. M.[1*], Isah, A.[2], Abdullahi, U.[3] and Sayuti, F. Y.[4]
[1, 2, 3 & 4]Department of Statistics, Federal University of Technology Minna, Nigeria.
*Corresponding Author Email: Oguche.samuel@futminna.edu.ng,
[2]aisah@futminna.edu.ng, [3]u.abdullahi@futminna.edu.ng , [4]fateesayuti555@gmail.com

Keywords:
Survival analysis, Human Immunodeficiency Virus, Cox proportional model, Gompertz model, Kaplan-Meier estimate.

ABSTRACT
This research utilised survival analysis to study and model data on human immunodeficiency virus (HIV) infection, among children aged one to ten years old. The study adopted the Gompertz parametric and Cox proportional semi parametric survival model to fit the HIV data and estimated the survival functions using the Kaplan-Meier estimator. The predictor variables of interest included age, weight, height, sex, mother-to-child transmission, residence, viral load, HIV and educational status of the parents. Purposive sampling technique was used to collect the required data from January, 2007 to December, 2021 at General Hospital Minna, Niger State. The analysis revealed that only viral load with a p-value of 0.000 was statistically significant predictor of the event, indicating it is significant at 0.05 level. This indicated lower risk of death and higher chance of surviving. The Gompertz model gave the best fit with Akaike Information Criterion of 569.17 and Bayesian Information Criterion of 609.19. Early initiation and consistent adherence to Antiretroviral Therapy (ART) can effectively suppress viral load. Children born to HIV-positive mothers, particularly those with unsuppressed viral loads, require regular monitoring due to their increased risk of death from mother-to-child transmission.

INTRODUCTION

Human Immunodeficiency Virus (HIV) remains one of the leading causes of morbidity and mortality among children worldwide, especially in sub-Saharan Africa, where vertical transmission from mother to child accounts for a large proportion of paediatric infections. Despite the availability of Antiretroviral Therapy (ART), many infected children face reduced life expectancies due to delayed diagnosis and treatment, as well as the complications associated with opportunistic infections (UNAIDS, 2022).

In 2023, there were approximately 95,000 children between the ages of 1 to 10 years living with HIV in Nigeria. The country continues to face challenges in addressing paediatric HIV due to high rates of mother-to-child transmission and limited access to treatment services. Children in this age group represent a significant portion of the estimated 140,000 children aged 14 years and below living with HIV in Nigeria (UNICEF, 2023).

When infected with HIV, children present with a wide spectrum of clinical presentations ranging from asymptomatic HIV infection to acquired immunodeficiency syndrome (AIDS), characterized by severe immunosuppression. Predominant presentations include respiratory infections, malnutrition, diarrhoea, and septicaemia.

Survival analysis is essential for predicting patients' time-to-event outcomes and aiding healthcare practitioners in making the best treatment decisions (Wang *et al*., 2017), not only in disease analysis or monitoring procedures but also in assisting with the quantitative and qualitative improvement of preventive medicine (lifestyle interventions, vaccine efficacy and screening programmes, among others). In addition to its use in healthcare, survival analysis plays a key role in decision-making across a variety of disciplines of management (Wang *et al.*, 2017). According to Lee and Wang (2003) and used by Suleman *et al.* (2025), both non-parametric methods like the Kaplan-Meier estimator and semi-parametric approaches such as the Cox proportional hazards model provide robust tools for analyzing censored survival data. Their work also emphasizes the importance of choosing appropriate parametric models like the Gompertz model when the hazard function follows a specific pattern, making it especially relevant for studies assessing long-term outcomes in chronic conditions such as HIV/AIDS. The Gompertz model, first introduced by Gompertz in

1825, is a widely utilized sigmoid model that is particularly effective in fitting growth data and other phenomena characterized by S-shaped curves. The model is named after Gompertz, who originally applied it to describe the growth of human populations (Chu, 2020). Cox's Proportional Hazards Modelling (PHM) is a partial likelihood perspective in which the baseline hazard rate is an unspecified nuisance function (Li, 2014). It can also be described as the basic modelling or technique used in exploring the relationship between the survival experiments and potential risk factors.

Several studies across Nigeria have documented the clinical patterns, prevalence, and presentation of paediatric HIV/AIDS, highlighting regional similarities and differences. At a tertiary hospital, Onankpa *et al.* (2008) reported diverse manifestations of paediatric HIV over a five-year period, reinforcing the need for early diagnosis and consistent follow-up. Similarly, Oniyangi *et al*. (2006) observed a wide range of clinical features among HIV-infected children in Abuja, indicating delayed diagnosis in many cases. In Gwagwalada, Okechukwu *et al*. (2008) found that most children presented with advanced disease symptoms, stressing the importance of early intervention. Ogunbosi *et al.* (2009) noted comparable trends in Ibadan, with late presentation being a common challenge. Meanwhile, Ejiofor *et al.* (2010) conducted a seroprevalence study in Awka, which revealed significant HIV antibody positivity among paediatric patients, emphasizing the continued risk of mother-to-child transmission. In a more localized study, Abdulkareem *et al*. (2023) assessed the prevalence of HIV/AIDS in Chanchaga Local Government, Niger State, between 2018 and 2022, contributing important regional data relevant to understanding the burden of disease in northern Nigeria.

Ikhelowa *et al*. (2019), on relating Cox Proportional Hazard model to evaluate the determinant factors of survival time, and predict the clinical progression of HIV/AIDs ailment employed secondary data obtained from the Antiretroviral Rehabilitation Unit of Central Hospital, Agbor, Delta State, Nigeria. The statistics were extracted from Regular Patient Medical Registration. They used a sample of 1000 HIV/AIDs patients who were followed for a minimum predetermined period of 11 years and 3 months. From the sample, 64.2 percent were female and 35.8 percent were male, 8.6 percent of the patients were reported dead; while 91.4 percent of patients were censored. The Cox regression result indicated that the survival time of the HIV/AIDs patients is significantly related to sex, ART, enrolment date, and current age.

Furthermore, a systematic review by Warszawski *et al.* (2017) synthesized evidence from multiple studies to assess the survival of vertically HIV-infected children in high-income countries. The review included data from 16 studies and highlighted the substantial improvements in survival rates over the years, attributed to advancements in ART and comprehensive care. The findings underscored the importance of access to quality healthcare services and ongoing monitoring to optimize survival outcomes in HIV-infected children.

A study by Njom *et al.* (2016) examined the survival of HIV-infected infants in Yaoundé, Cameroon. The researchers followed a cohort of 182 infants and assessed their survival outcomes over a period of 48 months. The study revealed that infants with late HIV diagnosis, low birth weight, and severe immunosuppression had significantly higher mortality rates. The findings underscored the need for timely HIV diagnosis, early initiation of ART, and nutritional interventions to enhance the survival prospects of HIV-infected infants.

This study explores survival outcomes for HIV-infected children aged one to ten years using survival analysis and fits a Gompertz parametric and Cox Proportional semi-parametric survival models to identify significant predictors of survival and also check for the model with the best fit for the data.

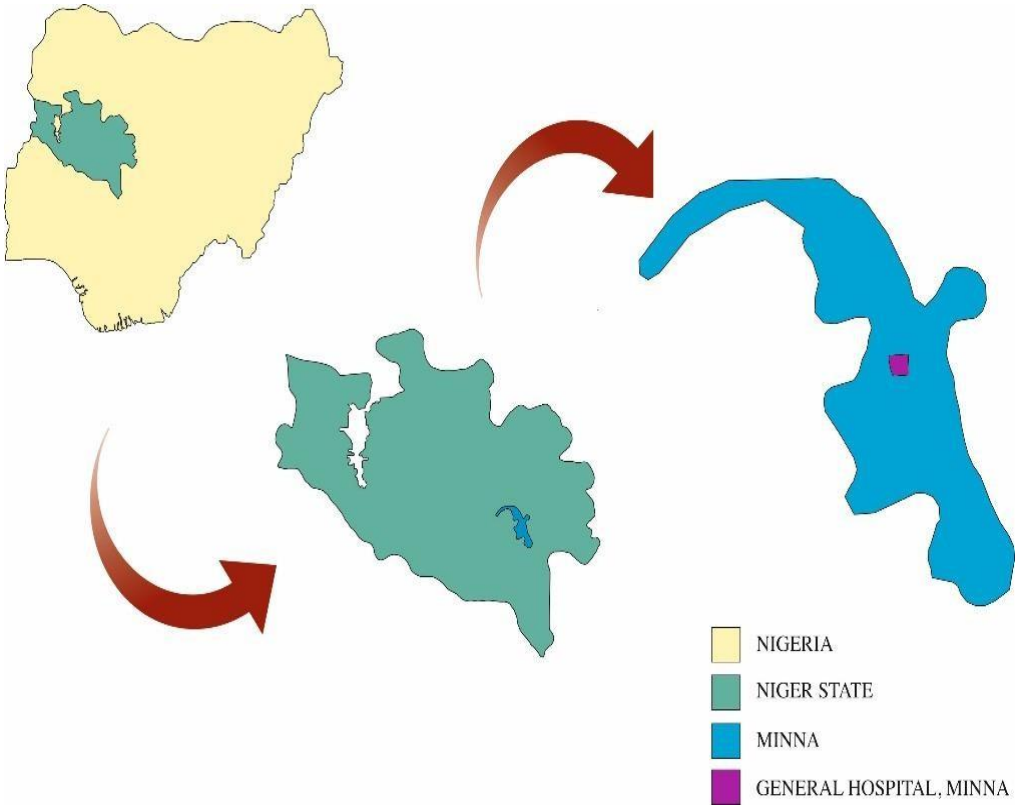

Figure 1: Map showing Minna General Hospital, Niger State, Nigeria
Source: Ministry of Land and Housing

MATERIALS AND METHODS

The data for this study was collected from the General Hospital Minna, Niger State, Nigeria. The hospital has a record of 1,021 registered children with HIV/AIDS cases that have been followed over a 15-year period; from January 2007 to December 2021.

From the population of 1,021 cases, a sample of 281 cases was selected for detailed analysis. Given the nature of the study, a purposive sampling method was employed. This sampling technique involves selecting cases based on specific characteristics relevant to the research objectives rather than choosing them randomly. A purposive sampling technique was used because survival analysis often requires cases with certain attributes, such as complete follow-up data or events (like survival or mortality) occurring within the study period.

Data were collected on patients, focusing on relevant socio-demographic and clinical variables. Patient records were carefully selected based on the following criteria: the individual was HIV-positive; was 10 years of age or younger; was registered at General Hospital Minna; was receiving treatment specifically at the Heart to Heart Unit of the hospital; and had been diagnosed with HIV as of January 2007. The collected data were entered into Microsoft Excel and subsequently analyzed using Microsoft Excel, STATA, and SPSS.

Gompertz Model

The Gompertz modelintroduced by B. Gompertz in 1825, is a widely utilized sigmoid model that is particularly effective in fitting growth data and other phenomena characterized by S-shaped curves. The model is named after Gompertz, who originally applied it to describe the growth of human populations (Chu, 2020).

The probability density function (pdf) of the Gompertz distribution with two shape parameters $\sigma$and z; $\sigma > 0$ is given by Collett*et al*. (2015) is given as

$$f(t \mid \sigma, z) = \sigma e^{zt} \exp^{-\frac{\sigma}{z}(e^{zt}-1)}, t > 0 \quad (1)$$

where σ captures the starting level of survival
z gives the rate of survival
t denotes the time considered for the study
Thus, following equation (1), the survival function is given by:

$$S(t) = \int_t^{\infty} \sigma e^{zt} e^{\frac{-\sigma(e^{zt}-1)}{z}} dt \quad (2)$$

$$= e^{\frac{-\delta(e^{\kappa t}-1)}{\kappa}} \quad (3)$$

and the hazard function is given by

$$h(t) = \frac{f(t)}{S(t)} = \sigma e^{zt} \quad (4)$$

Clearly, the hazard function of the Gompertz distribution increases with time when z<0, decreases

with time when z>0, and is constant when κ = 0. Hence, the Gompertz hazard function increases or decreases monotonically

The hazard function of the Gompertz distribution increases with time when $\kappa > 0$, decreases with time when $\kappa < 0$, and is constant when $\kappa = 0$ (Collett et al., 2015).

Applying the logarithm transformation to both sides of equation (4) yields (5)

$$\ln[h(t)] = \alpha + \kappa t \quad (5)$$

The hazard function $h(t) = \delta e^{zt}$ is an exponential function, which means the relationship between $h(t)$ and t is not linear but rather follows an exponential growth pattern. By taking the natural logarithm of both sides, you convert this exponential relationship into a linear one:

$$\ln[h(t)] = \ln\delta + zt \quad (6)$$

Here, $\ln\delta$ becomes a constant term (denoted as α), and t is the term with the variable t. This results in a linear equation:

$$\ln[h(t)] = \alpha + zt \quad (7)$$

where α denote ln δ. Hence, the Gompertz distribution is suitable for a given dataset where the hazard rate changes exponentially over time.

Cox's Proportional Hazards

Cox's Proportional Hazards Modeling (PHM) is a partial likelihood perspective in which the baseline hazard rate is an unspecified nuisance function (Liu, 2014). It can also be described as the basic modelling or technique used in exploring the relationship between the survival experiments and potential risk factors. Although, the model is based on the theory of hazards proportionality as such, no particular form of likelihood distribution is assumed for the survival times. The model is thus semi-parametric in nature and does not require a specified distribution for survival times.

The formula for the Cox proportional hazard model is:

$$h(t|X) = h(t)\, exp(X_1\beta_1 + \cdots + X_p\beta_p) \quad (8)$$

where:

$X_1,\ldots,X_p$ are the predictors, $\beta_1, \ldots, \beta_p$ are the coefficients associated with each predictor.

The predictors are assumed to act additively on log $h(t|X)$ and changes linearly with the βs. The effect of the predictors is assumed to be the same at all times (Collett *et al.*, 2015).

Kaplan-Meier estimate of the Survival and Hazard function

$$\hat{S}(t) = \prod_{j=1}^{k}\left(1 - \frac{d_j}{n_j}\right) = \prod_{j=1}^{k}\left(\frac{n_j - d_j}{n_j}\right) \quad (9)$$

For $t_k \le t < t_{k+1}, k = 1,2,\ldots,r$ with $\hat{S}(0) = 1$ and $\hat{S}(\infty) = 0$

where the product is taken over all j such that $k \le 1$. The estimator provides an estimate of the probability that a subject survives past time t, given the available data up to that time point.

Taking d_j as the number of events at the j^{th} event time denoted by *t(j)*, and nj the number of individuals at risk at time $t_{(j)}$, the hazard function at time t in the interval $\left[t_{(j)}, t_{(j+1)}\right)$ is estimated by

$$\hat{h}(t) = \frac{d_j}{n_j t_j} \quad (10)$$

$\hat{h}(t) = \frac{d_j}{n_j t_j}$ is known as the Kaplan-Meier estimate of the hazard function.

Validation Techniques

Two evaluation techniques are used to evaluate the performance of the methods.

Akaike Information Criterion

For model comparison and selection, the Akaike Information Criterion (AIC) is a flexible criterion. If the models were fitted to the same dataset, AIC can be used to compare both nested and non-nested models. AIC takes the form of a penalized likelihood, where the number of free parameters in the model determines the value of the "penalty" term. Three complementary perspectives, including information theory, prediction accuracy, and Bayesian statistics, can be used to understand the significance of AIC (Posada and Buckley, 2004). The number of free parameters in the model (K) and the maximal likelihood of a fitted model (L) are used to calculate the AIC as given below.

$$AIC = 2k - 2ln(\mathcal{L}) \quad (11)$$

Bayesian Information Criterion

The Bayesian Information Criterion (BIC) is a widely recognized generic model selection technique that prioritizes succinct models over complicated models by adding a penalty that is contingent on the number of factors being estimated in the model (Wang, 2017). One form for calculating the BIC is given by.

$$AIC = T_m - df_m ln(N) \quad (12)$$

where T_m represents the model's hypothesized chi-square statistic. In this form, a BIC less than 0 favours the hypothesized model, but a BIC greater than 0 favours the saturated model (i.e., the model that permits all observed variables to be intercorrelated with no assumed model structure). The BIC can also be used to evaluate two rival models.

RESULTS AND DISCUSSIONS

Table 1: Summary Statistics of HIV Data of children

Variables	Category	Count	Percent
Sex	Male	171	60.9
	Female	110	39.1
Mother-To-Child-Transmission	Yes	200	71.2
	No	81	28.8
Residence	Rural	76	27.0
	Urban	205	73.0
Present Viral Load	Unsuppressed	29	10.3
	Suppressed	252	89.7
Event	Dead	92	32.7
	Alive	189	67.3

Table 1 shows that 60.9% of children were male and 39.1% female, with 71.2% born HIV-positive due to Mother-To-Child Transmission. 73% lived in urban areas, and 89.7% had suppressed viral loads. 32.7% died, while 67.32% were alive.

Table 2: Kaplan Meir Estimate of the Median Survival for Viral Load

Category	N	Events	% Event	Median	Std Error	0.95LCL	0.95UCL
Suppressed	252	66	26	366	92.33	185.03	546.97
Unsuppressed	29	26	90	31	7.86	15.59	46.41
Overall	281	92	33	244	86.23	74.99	413.01
Chi-square	44.94	**Df**	1	**p-value**	0.000		

Table 2 shows a significant difference in children's viral load, with suppressed children having a 50% chance of survival up to 547 months, and unsuppressed children having a median survival time of 31 months.

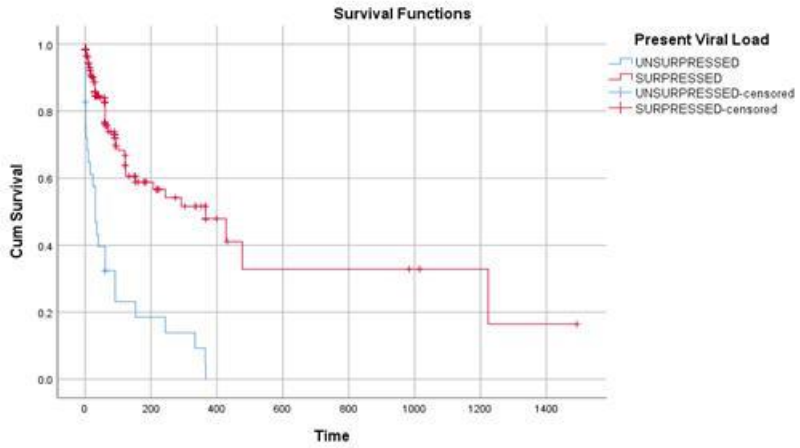

Figure 2: Estimate of Survival Curve based on Viral Load

Figure 2 shows that children with unsuppressed viral load have a higher survival probability at early age, but as time progresses, the probability decreases. Children with suppressed viral load have a higher survival probability, while those with unsuppressed viral load have a higher risk of death or negative health

outcomes. The curve's less steepness indicates more similar survival probabilities between groups.

Table 3: Gompertz Survival Model

Variables	Hazard Ratio	Std Error	Z	p-value	0.95 LCI	0.95 UCI
Age	1.022942	0.0502841	0.46	0.644	0.9289856	1.126401
Sex	1.187103	0.2764228	0.74	0.461	0.7521108	1.873677
Weight	.9808286	.0144076	-1.32	0.188	.9529929	1.009477
Height	1.008807	.0097318	0.91	0.363	.9899119	1.028062
MTCT	0.75783	1.163146	-0.18	0.857	.0374211	15.34714
HIV status	1.013582	1.530299	0.01	0.993	.052568	19.54325
Educational status	1.010017	.1499402	0.07	0.946	.755032	1.351115
Residence	.686904	.1626265	-1.59	0.113	.4318885	1.092497
Viral load	.2458283	.0674764	-5.11	0.000	.1435454	.4209927

Table 3 shows that viral load is the only significant variable, reducing HIV deaths in children by 2.458%. Factors such as age, sex, height, HIV status, and education increase death chances, while weight, MTCT, residence, and viral load decrease them.

Table 4: Cox Proportional Hazard Model

Variables	HR	Std Error.	Z	p-value	0.95LCL	0.95UCL
Age	1.026313	0.0494942	0.54	0.590	0.9337499	1.128053
Sex	1.226338	.2832994	0.88	0.377	.7797796	1.928627
Weight	0.9854807	.013505	-1.07	0.286	.9593637	1.012309
Height	1.006891	.0090547	0.76	0.445	.9892993	1.024795
MTCT	0.8424571	1.211077	-0.12	0.905	.0503384	17.116
HIV status	0.9851171	1.390364	-0.01	0.992	.0619623	15.66202
Edu status	1.044109	0.1528465	0.29	0.768	.7836806	1.391082
Residence	0.7004089	.1652219	-1.51	0.131	.441122	1.112102
Viral load	0.2329946	.0635108	-5.34	**0.000**	.1365593	.3975307

Table 4 shows that viral load significantly reduces HIV-related deaths in children by 0.232. Factors such as age, sex, height, HIV status, and education increase the chances of death, while weight, MTCT, residence, and viral load decrease the chances of death.

Akaike's and Bayesian Information Criterion of Cox Proportional Hazard Model

Table 5: Cox Proportional Hazard Model Akaike's and Bayesian Information Criterion

Observations	LL(null)	LL(model)	Df	AIC	BIC
281	-437.3433	-420.0448	9	858.0897	890.8348

Table 5 shows the Akaike's and Bayesian Information Criterion for the Cox proportional model. It shows that the observations are 281, the log-likelihood of the null model is -437.3433, the log-likelihood of the model is -420.0448, the Degrees of Freedom is 9, the AIC is 858.0897 while the BIC is 890.8348.

Table 6: Comparisons of Gompertz and Cox Proportional Hazard models using AIC and BIC.

Model	LL(null)	LL(model)	Df	AIC	BIC
Gompertz	-290.8435	-273.5835	9	569.167	609.1889
Cox Proportional Hazard	-437.3433	-420.0448	9	858.0897	890.8348

The tables 6 shows that the AIC for the Gompertz model (569.167) is lower than the AIC for the Cox proportional model (858.0897), which shows that the Gompertz model provides a better fit to the data. It shows that the BIC for the Gompertz model (609.1889) is lower than the BIC for the Cox proportional (890.8348) which shows that the Gompertz model provides a better fit to the data. These results shows that the Gompertz model provides a good fit to the data.

CONCLUSION

This study analyzed survival outcomes of 281 HIV-infected children at General Hospital Minna, revealing a 67.26% survival rate and a median survival time of 244 months. Survival probability declined with age, highlighting increased risks over time. Key predictors

of survival outcomes were identified, with viral load showing a statistically significant impact on mortality (p-value = 0.000). This indicated lower risk of death and higher chance of surviving. Factors such as age, sex, height, HIV status, and education level were insignificantly associated with an increased risk of death, whereas improvements in weight, Mother-To-Child Transmission (MTCT) rates, changes in residence, and reductions in viral load contributed to better survival outcomes. The Gompertz model provided a better fit than the Cox model, as shown by lower AIC and BIC values. Due to the purposive sampling method, the findings may not be generalizable beyond similar settings. The study emphasizes the importance of early intervention, viral load control, and addressing socio-demographic factors to improve survival among HIV-infected children.

REFERENCES

Abdulkareem, J. H., Abdulkarim, H., & Aliyu, F. Y. (2023). An assessment of prevalence of HIV/AIDS in Chanchaga Local Government from 2018 to 2022: Niger State, Nigeria. *American Journal of Clinical and Experimental Medicine, 11*(6), 108–113. https://doi.org/10.11648/j.ajcem.20231106.12

Chu, K. H. (2020). Fitting the Gompertz equation to asymmetric breakthrough curves. *Journal of environmental chemical engineering*, *8*(3), 103713.

Collett D., Staplin, N. D., Kimber, A. C., & Roderick, P. J. (2015). Dependent censoring in piecewise exponential survival models. *Statistical methods in medical research*, *24*(3), 325-341.

Ejiofor, O. S., Onyire, N. B., Chapp-Jumbo, A. U., Ofomata, J. A., Akpati, R. N., and Ezeokoli, N. O. (2010). The Seroprevalence of HIV antibodies in children attending Amaku General Hospital Awka, South-East Nigeria. *European J of Scientific Research*, *43*(2), 351-356.

Ikhelowa, A. P., Joseph, B. O., and Godstime, U. O. M. A. (2019). Survival Analysis of HIV/AIDs Patients under Antiretroviral Treatment at Central Hospital, Agbor, Delta State, Nigeria.

Li, M., Hu, Q., and Liu, J. (2014). Proportional hazard modeling for hierarchical systems with multi-level information aggregation. *IIE Transactions*, *46*(2), 149-163.

Njom Nlend, A. E., Motaze, A. C. N., Sandie, A., and Fokam, J. (2018). HIV-1 transmission and survival according to feeding options in infants born to HIV-infected women in Yaoundé, Cameroon. *BMC pediatrics*, *18*, 1-8.

Ogunbosi, B. O., Oladokun, R., Brown, B. J., and Osinusi, K. (2009). *HIV infection in children presenting at the University College Hospital, Ibadan*. Paper presented at the Paediatric Association of Nigeria Conference (PANCONF), Ibadan, Nigeria.

Okechukwu, A. A., Gambo, D., and Okechukwu, O. I. (2008). The clinical features of paediatric HIV/AIDS at presentation at the University of Abuja Teaching Hospital, Gwagwalada. *Nigerian Journal of Medicine, 17*(4), 433–438.

Onankpa, B., Airede, L., Paul, I., and Dorcas, I. (2008). Pattern of pediatric HIV/AIDS: A five-year experience in a tertiary hospital. *Journal of the National Medical Association, 100*(7), 821–825.

Oniyangi, O., Awani, B., & Iregbu, K. C. (2006). The pattern of paediatric HIV/AIDS as seen at the National Hospital Abuja, Nigeria. *Nigerian Journal of Clinical Practice, 9*(2), 153–158.

Suleman, I., Zakariyau, N. R., Oyegoke, O. A., Yahya, W. B., Amiru, F. M., and Umar, M. A. (2025). Notes on a modified exponential-gamma distribution: Its properties and applications. *Journal of Basics and Applied Sciences Research*, *3*(2), 12-26. https://dx.doi.org/10.4314/jobasr.v3i2.2

UNAIDS. (2022). AIDS epidemic update. Geneva: UNAIDS.

UNICEF. (2023). 2023 Global Snapshot on HIV and AIDS Progress and priorities for children, adolescents and pregnant women. Available at: https://data.unicef.org/country/cmr/

Wang, P., Li, Y. and Reddy, C.K. (2017). Machine Learning for Survival Analysis: A Survey. Aug. [Online]. Available: http://arxiv.org/abs/1708.04649

Wang, P., Li, Y., and Reddy, C. K. (2017). Machine learning for survival analysis: A survey. ACM Computing Surveys, 51(6). https://doi.org/10.1145/3214306

Warszawski, J., Le Chenadec, J., Faye, A., Dollfus, C., Mellul, S., Tréluyer, J. M., and Blanche, S. (2017). Long-term nonprogression of HIV infection in children: evaluation of the ANRS EPF-CO10 paediatric cohort. Antiviral Therapy, 22(2), 145