Science World Journal Vol. 21(No 2) 2026 https://dx.doi.org/10.4314/swj.v21i2.7
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

BAYESIAN INFERENCE FOR HUMAN IMMUNODEFICIENCY VIRUS (HIV) CLINICAL PROGRESSION: MODELLING CURRENT CLUSTER OF DIFFERENTIATION 4 (CD4) WITH PATIENT-LEVEL COVARIATES

Ogundeji K. Rotimi, *Chinenye C. Nkem-Nnake

Department of Statistics, University of Lagos, Akoka, Lagos, Nigeria

*Corresponding Author Email Address: chinenyenzenwata@gmail.com

ABSTRACT
This study aimed to model current CD4 lymphocyte counts among people living with HIV by examining the effects of age, sex, baseline CD4 levels, and changes in body weight. A Bayesian regression framework was applied to clinical data to incorporate prior information and quantify uncertainty in parameter estimates. The findings indicate that baseline CD4 count was the strongest predictor of current CD4 levels. At the same time, age showed a moderate positive association, and the effects of sex and short-term weight change were smaller and less certain. Overall, the results demonstrate that Bayesian regression provides a robust and informative approach for understanding CD4 dynamics and supporting clinical decision-making in resource-limited settings.

Keywords: HIV/AIDS; CD4 count; Bayesian regression; Disease progression; Nutritional status; Sub-Saharan Africa

INTRODUCTION

Human Immunodeficiency Virus (HIV) infection remains a major global public health challenge more than four decades after its discovery. According to the Joint United Nations Programme on HIV/AIDS (UNAIDS), approximately 39 million people were living with HIV worldwide in 2022, with sub-Saharan Africa bearing nearly two-thirds of the global burden (UNAIDS, 2023). Although antiretroviral therapy (ART) has transformed HIV into a manageable chronic condition, effective long-term management depends on accurate monitoring of immunological status and an improved understanding of disease progression at the individual level (Tang & Kaslow, 2012; Ford et al., 2017)

CD4 T-lymphocyte count is a central marker of immune function in HIV-infected individuals. Progressive depletion of CD4 cells in untreated infection leads to immune suppression and increased susceptibility to opportunistic infections (WHO, 2016; Ford *et al*., 2017). Even in the era of widespread ART use, CD4 counts remain clinically important, particularly in low- and middle-income countries where access to routine viral load testing is limited (Siedner *et a*l., 2016; Ford *et al*., 2017). Monitoring CD4 dynamics, therefore, continues to play a key role in assessing disease stage, guiding treatment decisions, and identifying patients at risk of poor outcomes (WHO, 2016).

Previous studies have shown that CD4 trajectories are influenced by multiple patient-level factors, including baseline CD4 count, age, sex, and nutritional status (Hughes *et al*., 2012; Filteau *et al*., 2015; Koethe & Heimburger, 2015). Younger individuals generally experience stronger immune recovery, while older patients may show slower or incomplete reconstitution (Hughes *et al*., 2012). Sex-related differences in immune response and the interaction between nutritional status and immune function further highlight the multifactorial nature of HIV progression (Hughes *et al*., 2012; Filteau *et al*., 2015). Accurately capturing these influences requires statistical methods that can handle biological variability and uncertainty (Gelman & Hill, 2007; Gelman *et al*., 2013).

Traditional frequentist regression approaches have been widely applied in HIV research, but often rely on restrictive assumptions and provide limited representation of uncertainty. Bayesian statistical methods offer a probabilistic framework that allows the incorporation of prior knowledge and the direct quantification of uncertainty through posterior distributions (Carlin & Louis, 2009; Betancourt, 2017; Carpenter *et al*., 2017; Oyelami & Ogundeji, 2025). This approach is particularly well suited to clinical data, where variability is high and prior biological understanding is well established (Gelman *et al*., 2013; Gabry *et al*., 2019).

The goal of this study is to apply a Bayesian regression model to investigate factors associated with current CD4 counts among people living with HIV. Specifically, the study examines the effects of age, sex, baseline CD4 level, and changes in body weight on current immunological status. By adopting a Bayesian framework, the study aims to provide more informative and clinically interpretable estimates of CD4 dynamics, thereby improving understanding of HIV disease progression and supporting evidence-based care in resource-limited settings.

MATERIALS AND METHODS

Study Design and Data

This study employed a secondary analysis of patient-level HIV monitoring data comprising demographic and clinical information for individuals diagnosed with HIV. The dataset consisted of 701 observations, with no missing values across all variables considered. The analysis was restricted to infected individuals, as all records had HIV status equal to 1.

Let $i = 1, \ldots, n$ index patients, where $n = 701$. The key variables extracted from the dataset were:

Age			(Age_i),
Sex			(Sex_i),
Initial	CD4	count	$(IntCD4_i)$,
Current	CD4	count	$(CntCD4_i)$,
Starting	body	weight	(SW_i),
Current	body	weight	(CW_i).

From these measurements, derived covariates were constructed as:

$$\Delta Weight_i = CW_i - SW_i \quad (1)$$

$\Delta CD4_i = CntCD4_i - IntCD4_i$ (2)

Sex was encoded as a binary indicator variable (e.g., 1 = female, 0 = male).

The primary outcome variable for analysis was the current CD4 count ($CntCD4_i$), a widely used biomarker for assessing immune function, HIV disease progression, and treatment response. The predictor variables included age, sex, baseline (initial) CD4 count, and change in body weight. CD4 count remains clinically important for monitoring treatment effectiveness and immunological recovery, particularly in settings where viral load testing may be limited (WHO, 2016; Ford *et al.*, 2017). The predictor variables included age, sex, baseline (initial) CD4 count, and change in body weight. These covariates were selected based on established clinical relevance and empirical evidence demonstrating their association with immune reconstitution and disease progression in HIV-infected populations (Hughes *et al.*, 2012; Filteau *et al.*, 2015; Koethe & Heimburger, 2015).

Bayesian Regression Framework

The objective of this study is to model current CD4 counts using a Bayesian regression framework. Let $i = 1, \ldots, n$ index patients in the study.

The response variable is defined as:

$Y_i = Current\ CD4\ for\ patient\ i$ (3)

The corresponding covariate vector is given by:

$X_i = (1, Age_i, Female_i, IntCD4_i, \Delta Weight t_i)$ (4)

where the first element represents the intercept term.

Likelihood Specification

We assume a Gaussian likelihood for the observed CD4 counts:Y_i

$|\beta, \sigma^2 \sim N({X_i}^T \beta, \sigma^2), i = 1, \ldots n,$ (5)

where

$\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)^T$is the vector of regression coefficients, and

$\text{squared notes}$the residual variance.

Equivalently, the linear predictor can be written explicitly as:

$\mu_i = \beta_0 + \beta_1 Age_i + \beta_2 Female_i + \beta_{3I} IntCD4_i + \beta_4 \Delta Weight_i$ (6)

Thus, the sampling model becomes:

$Y_i | \mu_i, \sigma^2 \sim N(\mu_i, \sigma^2).$ (7)

Combining equations (5)– (7), the Bayesian regression model specifies current CD4 count as a linear function of age, sex, baseline CD4, and weight change, with uncertainty captured by the variance parameter σ. This formulation provides a flexible probabilistic foundation for incorporating prior information and quantifying uncertainty in HIV disease progression.

Priors

To regularize estimation while maintaining biological plausibility, weakly informative priors were assigned to all model parameters (Ogundeji & Adeleke, 2020).

Each regression coefficient was assumed to follow a normal prior distribution:

$\beta_j \sim N(0, 10^2), j = 0,1,2,3,4,$ (8)

where β_0 denotes the intercept and $\beta_1, \ldots \beta_4$ correspond to the covariate effects.

The residual standard deviation was assigned a half-normal prior:

$\sigma \sim Half\ Normal\ (0,10)$ (9)

which constrains σ to be positive while remaining weakly informative.

Posterior Distribution

Let **Y**=$(Y_1, \ldots Y_n)^T$ denote the vector of observed current CD4 counts and **X** = $(X_1, \ldots X_n)^T$the corresponding design matrix. By Bayes' theorem, the joint posterior distribution of the parameters is given by:

$p(\beta, \sigma \mid \mathbf{Y}, \mathbf{X}) \propto p(\mathbf{Y} \mid \mathbf{X}, \beta, \sigma) p(\beta)\ p(\sigma),$ (10)

which can be written explicitly as

$$p(\beta, \sigma \mid Y, X) \propto \left[\prod_{i=1}^{n} N(Y_i \mid {X_i}^T \beta, \sigma^2)\right] \left[\prod_{j=1}^{4} N(\beta_j \mid 0, 10^2)\right] \quad (11)$$

Posterior Inference and Model Evaluation

Posterior inference was carried out using Hamiltonian Monte Carlo (HMC) sampling as implemented in the PyMC framework. Model convergence was evaluated using the potential scale reduction factor ($\hat{R}$), trace plots, and effective sample sizes. Posterior summaries, including posterior means, medians, and 95% credible intervals, were reported for all parameters.

Model adequacy was assessed through posterior predictive checks, comparing replicated data generated from the posterior predictive distribution with the observed CD4 counts to evaluate goodness of fit.

Bayesian Hierarchical Model: Extension and Posterior Factorization

The hierarchical structure of the model induces partial pooling, whereby information is shared across clinics through a common prior distribution on the random effects. This approach allows clinics with limited observations to borrow strength from the overall population means, while clinics with larger sample sizes contribute more strongly to their own estimates. Such shrinkage estimation improves parameter stability and reduces estimation variance, particularly in datasets with unbalanced group sizes (Gelman & Hill, 2007; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012)

From a Bayesian perspective, hierarchical modelling provides a principled mechanism for incorporating prior distributions on variance components and random effects, thereby improving inference under uncertainty (Gelman, 2006; McElreath, 2020). The use of weakly informative priors on variance parameters further regularizes estimation and prevents overfitting, especially in clinical datasets where subgroup sample sizes may vary considerably.

Consequently, the multilevel framework not only accounts for unobserved heterogeneity across clinics or time periods but also enhances predictive performance and interpretability of both group-level and individual-level effects.

. Suppose patients are nested within clinics indexed by $j = 1, \ldots, J,\ with\ i = 1, \ldots, n_j$ patients in clinic j.

Hierarchical Likelihood

Let Y_{ij} denote the current CD4 count for patient i in clinic j, and let X_{ij}denote the corresponding covariate vector. The hierarchical sampling model is specified as:

$Y_{ij} | \beta, u_j, \sigma^2 \sim N({X_{ij}}^T \beta + u_J, \sigma^2),\ i = 1, \ldots, n_j, j = 1, \ldots, J.$ (12)

Here, u_j represents a **clinic-specific random effect** capturing unobserved heterogeneity between clinics.

Hierarchical Priors

The clinic-level random effects are assumed to follow a normal distribution:

$$u_j \sim N(0, \tau^2), j = 1, \ldots, J, \quad (13)$$

Whesquared is the between-clinic variance.

Regression coefficients retain their weakly informative priors:

$$\beta_k \sim N(0, 10^2), k = 0,1,2,3,4, \quad (14)$$

Moreover, the variance parameters are assigned half-normal priors:

$$\sigma \sim Half - Normal\,(0,10),\ \tau \sim Half - Normal(0,10) \quad (15)$$

Posterior Factorization

Let. $u = (u_1, \ldots, u_J)^T$. By Bayes' theorem, the joint posterior distribution factorizes as:

$$p(\beta, \sigma, \tau, u\,|Y, X)\alpha\ p\,(Y|X, \beta, U, \sigma)\ p\,(\beta)\ p(\sigma)\ p\,(\tau) \prod_{j=1}^{J} p\,(u_j|\,\tau). \quad (16)$$

The hierarchical structure of the model induces partial pooling, whereby information is shared across clinics through a common prior distribution on the random effects. This modeling strategy allows clinics with limited observations to borrow strength from the overall population means, while clinics with larger sample sizes contribute more strongly to their own estimates. Consequently, the approach improves parameter stability and reduces estimation variance, particularly in datasets with uneven group sizes. Partial pooling also provides a principled way to account for unobserved heterogeneity across clinics or regions, thereby producing more reliable group- and individual-level inferences (Gelman & Hill, 2007; Gelman *et al.*, 2013; McElreath, 2020).

Computational Implementation

All statistical analyses were conducted using Python version 3.11. All Bayesian models were implemented in the probabilistic programming framework PyMC (Salvatier, Wiecki, and Fonnesbeck, 2016). Posterior inference was performed using Hamiltonian Monte Carlo sampling, which efficiently explores high-dimensional posterior distributions and improves convergence behavior in complex hierarchical models (Betancourt, 2017; Carpenter *et al.*, 2017; Ogundeji & Adeleke, 2022).

For each model, four independent Markov chains were run, with 2,000 iterations per chain. The first 1,000 iterations of each chain were discarded as warm-up (burn-in) to reduce dependence on initial parameter values and improve sampling stability. Model convergence was evaluated using standard Bayesian diagnostic criteria, including the potential scale reduction factor ($\hat{R}$), trace plots, and effective sample size estimates (Gelman *et al.*, 2013; Vehtari et al., 2017; McElreath, 2020).

Model adequacy was further examined using posterior predictive checks, which compare simulated data from the posterior predictive distribution with observed CD4 counts. This procedure helps detect systematic model misfit and ensures that the assumed likelihood structure adequately represents the empirical data (Spiegelhalter *et al.*, 2002; Gabry *et al.*, 2019).

Results are presented using posterior summaries, including posterior means, medians, and 95% credible intervals, to provide a comprehensive description of parameter uncertainty and facilitate clinical interpretation. Tables and figures were used to enhance clarity and support visual assessment of model behaviour and inference patterns.

RESULTS

Posterior Estimates

The Bayesian regression model demonstrated satisfactory convergence, with $\hat{R}$ values of approximately 1.00 for all parameters and effective sample sizes exceeding 1000. Posterior summaries of the regression coefficients are presented in **Table 1**.

Table 1: Posterior summaries for predictors of Current CD4 count

Parameter	Posterior Mean	95% CrI Lower	95% CrI Upper
Intercept	79.80	–	–
Age	8.06	5.61	10.50
Female	8.56	-0.39	17.50
Initial CD4	1.22	0.98	1.46
ΔWeight	1.05	-1.26	3.35
Residual SD	56.65	–	–

Baseline CD4 count exhibited the strongest positive association with current CD4 levels, with its 95% credible interval excluding zero. Age also showed a positive association, whereas the effects of sex and weight change were uncertain, as their credible intervals included zero. The residual standard deviation indicates substantial unexplained variability beyond the included predictors.

Posterior Predictive Checks

Posterior predictive checks indicated adequate model fit. The posterior predictive p-value for the mean was 0.480, and for the variance was 0.477, both close to 0.50, suggesting that the model reproduces the central tendency and variability of the observed CD4 distribution without systematic bias (Table 2).

Table 2. Posterior predictive p-values

Discrepancy	Posterior Predictive p-value
Mean	0.480
Variance	0.477

Figure 1 presents the posterior predictive distribution compared with the observed data. The model captures the central peak of the

Science World Journal Vol. 21(No 2) 2026
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

distribution; however, the observed data exhibit heavier tails than the replicated data, suggesting some underestimation of extreme CD4 values.

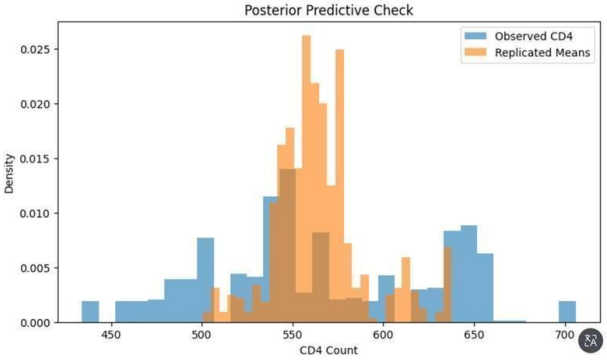

Figure 1. Posterior predictive distribution of current CD4 counts compared with observed data.

Hierarchical Model Results

Results from the hierarchical Bayesian model are summarized in Tables 3 and 4. Standardized population-level slopes are shown in Table 3.

Table 3. Hierarchical model: standardized population slopes

Predictor:	Age	Female	Initial CD4	ΔWeight
Estimate:	15.94	6.56	20.26	-3.36

Initial CD4 remained the dominant predictor under partial pooling, followed by age, while the effects of sex and weight change were smaller. Group-level intercepts reflecting partial pooling across years are shown in Table 4.

Table 4. Partially pooled year-level intercepts

Group (Year)	**Intercept**
Base (4)	550.02
Year 5	556.95
Year 6	587.77

Later years showed higher adjusted intercepts relative to the reference group, suggesting temporal shifts in expected CD4 levels after accounting for covariates.

DISCUSSION

The primary objective of this study was to identify determinants of current CD4 count among people living with HIV (PLHIV) using a Bayesian regression framework that explicitly accounts for uncertainty in clinical inference. The results demonstrated that baseline immunological status, particularly the initial CD4 count, was the strongest predictor of the current CD4 count. This finding is biologically plausible and consistent with established HIV treatment literature emphasizing the importance of early diagnosis and prompt initiation of antiretroviral therapy (ART). Similar public health evidence has been promoted by the Joint United Nations Programme on HIV/AIDS, which advocates test-and-treat strategies to reduce late presentation and improve long-term clinical outcomes.

The dominant influence of baseline CD4 count suggests that immune recovery is strongly dependent on disease stage at treatment initiation. Patients who begin treatment with higher CD4 levels are more likely to maintain superior immunological status over time, supporting policies that encourage early HIV screening and linkage to care. This observation is also consistent with World Health Organization treatment guidelines, which recommend early initiation of ART to preserve immune function and reduce the risk of opportunistic infections.

Short-term weight change showed a modest positive association with current CD4 count, highlighting the role of nutritional and metabolic status in immune recovery. Improved nutritional condition is associated with better treatment response and reduced mortality among PLHIV, particularly in resource-limited settings. Weight loss in HIV patients may reflect disease severity, socioeconomic vulnerability, or concurrent infections, all of which can impair immune reconstitution even under viral suppression. These findings align with clinical nutrition studies emphasizing integrated HIV management strategies, including dietary assessment and counseling (Gelman & Hill, 2007; Filteau *et al.*, 2015; Koethe & Heimburger, 2015).

Age showed a small positive association with the current CD4 count, although the effect size was relatively small. This result differs slightly from some reports indicating slower immune recovery among older patients. One possible explanation is that older individuals in the cohort may have demonstrated better treatment adherence or more consistent clinic attendance, potentially offsetting biological effects associated with immunosenescence. Variations in population structure, timing of ART initiation, and healthcare accessibility may also explain differences from previous studies (Hughes *et al.*, 2012; Siedner *et al.*, 2016).

The weak and statistically imprecise association observed for female sex suggests substantial heterogeneity in immune recovery between sexes. Previous studies have reported inconsistent findings on sex-specific CD4 responses, with some showing faster immune recovery in females, while others find minimal differences after accounting for baseline disease severity and treatment timing (Tang & Kaslow, 2012). The uncertainty around this parameter in the Bayesian model is therefore informative, suggesting that sex may not be a major predictor of immunological outcome compared with baseline disease status.

From a methodological perspective, the Bayesian multilevel modelling framework contributes to HIV clinical outcome research by providing probabilistic inference rather than relying on dichotomous significance testing. Posterior predictive checks showed good model adequacy, with Bayesian p-values close to 0.5, indicating that the likelihood structure captured both central tendency and dispersion of CD4 count distributions (Spiegelhalter *et al.*, 2002). The use of weakly informative priors helped stabilize parameter estimation while allowing empirical data to dominate posterior inference, thereby reducing risk (Carlin & Louis, 2009; Gelman *et al.*, 2013).

The hierarchical partial pooling structure improved estimation

reliability by shrinking noisy subgroup intercepts toward the population mean. This approach is particularly advantageous in clinical datasets with unequal clinic or regional sample sizes because it reduces variance inflation and improves predictive stability (Gelman & Hill, 2007; McElreath, 2020). The observation that later programmatic groups exhibited higher adjusted intercepts may reflect gradual improvements in HIV service delivery, including earlier ART initiation, enhanced adherence monitoring, and strengthened primary healthcare systems.

Some results deviated slightly from initial expectations, particularly the relatively small effect sizes observed for demographic variables such as age and sex. These differences may be attributable to residual confounding, limited heterogeneity within the study population, or the omission of clinically important predictors such as viral load suppression status, ART regimen type, or opportunistic infection history. Since the study design was cross-sectional, causal inference regarding covariate effects on CD4 dynamics is not appropriate. Longitudinal modelling would be more suitable for capturing the immune trajectory and treatment response over time.

Overall, the study demonstrates that Bayesian hierarchical modelling is a robust approach for HIV clinical outcomes. The framework enhances clinical interpretability by quantifying parameter uncertainty and supporting probabilistic decision-making. The findings reinforce the importance of early HIV detection, nutritional support, and sustained treatment engagement in promoting immune recovery. Future research should extend this work to longitudinal CD4 trajectory modelling, incorporate viral load measurements, and evaluate treatment regimen effects to support personalized HIV care in resource-limited settings.

Conclusion

This study applied a Bayesian regression modelling framework to examine the determinants of current CD4 count among people living with HIV, incorporating age, sex, baseline CD4 level, and changes in body weight as key predictors. The findings consistently demonstrate that baseline CD4 count is the most important predictor of current immunological status, underscoring the critical role of early diagnosis and timely initiation of antiretroviral therapy in preserving immune function.

Age showed a positive but modest association with current CD4 levels, whereas sex and short-term weight change showed weaker and more uncertain associations. These results suggest that demographic factors alone explain less of the variation in CD4 than baseline disease status, highlighting the dominant influence of initial immune condition on subsequent recovery trajectories.

The Bayesian modelling framework proved particularly valuable in this context by allowing direct quantification of uncertainty through posterior distributions and credible intervals, rather than relying solely on point estimates. Posterior predictive checks further indicated that the model adequately captured the central tendencies and variability of CD4 counts, supporting its suitability for clinical data analysis.

The hierarchical extension of the model improved interpretability and stability through partial pooling, enabling more reliable estimation of group-level effects and accounting for unobserved heterogeneity across clusters. This strengthens the approach's applicability in real-world healthcare settings, where data are often unevenly distributed.

Overall, the study demonstrates that Bayesian methods provide a robust and clinically meaningful framework for understanding HIV disease progression. The results reinforce the importance of early ART initiation, continuous monitoring of immunological status, and integrated care approaches that consider both clinical and nutritional factors. Future work should focus on longitudinal modelling of CD4 trajectories and the inclusion of additional biomarkers such as viral load and treatment adherence to further improve predictive accuracy and support personalized HIV care strategies in resource-limited settings.

REFERENCES

Abubakar, M.S. (2006). 'A new disease of dogs from Kaduna', *Science World Journal*, pp. 23–39. Available at: http://www.scienceworldjournal.com.

Betancourt, M. (2017). 'A conceptual introduction to Hamiltonian Monte Carlo', *arXiv preprint arXiv:1701.02434*.

Carlin, B.P. & Louis, T.A. (2009). *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd edn. Boca Raton: Chapman & Hall/CRC.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). 'Stan: A probabilistic programming language', *Journal of Statistical Software*, 76, pp. 1–32.

Filteau, S., PrayGod, G. & Kasonka, L. (2015). 'Nutritional status and response to antiretroviral therapy in adults with HIV infection', *Proceedings of the Nutrition Society*, 74, pp. 1–10.

Ford, N., Meintjes, G., Pozniak, A., Bygrave, H., Hill, A., Peter, T. & Vitoria, M. (2017). 'The future role of CD4 cell count for monitoring antiretroviral therapy', *The Lancet Infectious Diseases*, 17, pp. e76–e82.

Friis, H. & Michaelsen, K.F. (1998). 'Micronutrients and HIV infection: A review', *European Journal of Clinical Nutrition*, 52, pp. 157–163.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). 'Visualization in Bayesian workflow', *Journal of the Royal Statistical Society: Series A*, 182, pp. 389–402.

Gelman, A. (2006). 'Prior distributions for variance parameters in hierarchical models', *Bayesian Analysis*, 1(3), pp. 515–533.

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2013). *Bayesian Data Analysis*. 3rd edn. Boca Raton: Chapman & Hall/CRC.

Hughes, M.D., Ribaudo, H.J., & Bosch, R.J. (2012). 'Sex and age differences in immune reconstitution after initiating combination antiretroviral therapy', *Clinical Infectious Diseases*, 54, pp. 854–862.

Koethe, J.R. & Heimburger, D.C. (2015). 'Nutritional aspects of HIV-associated wasting in sub-Saharan Africa',

American Journal of Clinical Nutrition, 102, pp. 1–8.

Little, R.J.A. and Rubin, D.B. (2002) *Statistical Analysis with Missing Data*. 2nd edn. New York: Wiley.

McElreath, R. (2020). *Statistical Rethinking*. 2nd edn. Boca Raton: Chapman & Hall/CRC.

Mugavero, M.J., Amico, K.R., Horn, T., & Thompson, M.A. (2013). 'The state of engagement in HIV care in the United States: From cascade to continuum to control', *Clinical Infectious Diseases*, 57, pp. 1164–1171.

Ogundeji, R.K. & Adeleke, I.A. (2020). 'Derivation of prior and posterior distributions of regression parameters based on Zellner's g-priors', *FUW Trends in Science & Technology Journal*, 5(3), pp. 821–826.

Ogundeji, R.K. & Adeleke, I.A. (2022). 'Zellner's g-priors in Bayesian model averaging of large model space using Markov chain Monte Carlo model composition applicable under Bayesian model sampling', *BIMA Journal of Science and Technology*, 6(2), pp. 7–17.

Oyelami, A. & Ogundeji, R. (2025). 'Bayesian seasonal autoregressive integrated moving average: Modelling two decades of inflation dynamics in Nigeria', *Science World Journal*, 20(1), pp. 56–63.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd edn. Thousand Oaks: Sage.

Salvatier, J., Wiecki, T.V., & Fonnesbeck, C. (2016). 'Probabilistic programming in Python using PyMC3', *PeerJ Computer Science*, 2, e55.

Siedner, M.J., Ng, C.K., Bassett, I.V., Katz, I.T., Bangsberg, D.R., & Tsai, A.C. (2016). 'Trends in CD4 count at presentation to care and treatment initiation in sub-Saharan Africa, 2002–2013: A meta-analysis', *Clinical Infectious Diseases*, 62, pp. 377–387.

Snijders, T.A.B. & Bosker, R.J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2nd edn. London: Sage.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. (2002). 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society: Series B*, 64, pp. 583–639.

Tang, J. & Kaslow, R.A. (2012). 'The impact of host genetics on HIV infection and disease progression in the era of highly active antiretroviral therapy', *AIDS*, 26, pp. 1233–1242.

UNAIDS (2023) *UNAIDS Data 2023*. Geneva: Joint United Nations Programme on HIV/AIDS.

Vehtari, A., Gelman, A., & Gabry, J. (2017). 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC', *Statistics and Computing*, 27, pp. 1413–1432.

Vehtari, A., Simpson, D., Gelman, A., Yao, Y. & Gabry, J. (2021). 'Ranking Markov chain Monte Carlo draws by Pareto smoothed importance sampling', *Statistics and Computing*, 31, p. 24.

World Health Organization (2016) *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection*. Geneva: World Health Organization.