Pakistan
Journal of Biological Sciences

Research Article
Effects of Antiretroviral Therapy on CD4+ Cell Count, HIV Viral Load and Death in a South African Cohort: A Modelling Study

[1,3]Claris Shoko, [1]Delson Chikobvu and [2]Pascal O. Bessong

[1]Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, P.O. Box 339, 9300 Bloemfontein, South Africa
[2]HIV/AIDS and Global Health Research Programme, University of Venda, PMB X5050 0950, Thohoyandou, South Africa
[3]Great Zimbabwe University, P.O. Box 1235, Masvingo, Zimbabwe

Abstract

Background and Objective: Combination antiretroviral therapy (cART) has improved the survival of HIV infected patients significantly. However, in some patients, survival is not guaranteed due to several factors that are either individual-based or cART based. This study presents an HIV, AIDS, Death (HAD) model to analyse the survival of patients on cART. **Materials and Methods:** Continuous-time Markov models are fitted based on the states occupied for an HIV, AIDS and Death (HAD) model. These states are based on CD4 cell count. Factors that affect the survival of HIV-infected patients on cART are also analyzed. These, among others, include age, gender, routinely collected viral load, time on treatment, non-adherence and peripheral neuropathy. **Results:** Patients with higher viral loads than expected are 11.1 times more likely to be at risk of HIV progression to the AIDS state and 1.1 times more likely to be at risk of mortality from a CD4 cell count state above 200 cell/mm^{3}compared to patients with lower viral loads. Non-adherence to treatment increases the risk of transition from CD4 cell count state above 200 cell/mm^3 to the AIDS state by 2.2 folds. Patients who were non-adherent to treatment are 3.8 times more likely to transit from the CD4 state above 200 cell/mm^3 to death compared to patients who were adherent to treatment. Patients are expected to recover from the AIDS state after one year of treatment. **Conclusions:** Recovery from AIDS state by HIV infected patients on cART is likely to occur after one year of cART treatment. However, if the viral load remains higher than expected, this increases risks of immune deterioration even after having achieved normal CD4 cell counts and consequently, mortality risks are increased.

Key words: HAD model, Kolmogorov's backward equations, antiretroviral therapy, multistate modelling, markov jump process

Corresponding Author: Claris Shoko, Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, P.O. Box 339, 9300 Bloemfontein, South Africa Tel: 00263772587610

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

The availability of new antiretroviral drugs and highly active antiretroviral therapy (HAART) has led to a significant reduction in HIV morbidity and mortality[1]. The pharmacological activity of HAART has inhibitory effects on human immunodeficiency virus (HIV) duplication and has shown a significant reduction in acquired immunodeficiency syndrome (AIDS) epidemics as well as deaths[2]. Inhibitory of HIV duplication is generally associated with a steady increase in the CD4 cell count and results in improved clinical outcomes[3]. This increases the chances of survival for HIV infected patients.

Sieleunou *et al.*[3] used a retrospective cohort study of HIV patients older than 15 years in a rural centre in the Far-North Province of Cameroon to explore determinants of survival of HIV patients on antiretroviral therapy (ART). They used Kaplan-Meier analysis to estimate survival and Cox proportional hazard models to explain survival. Their findings show that CD4 cell count, gender and clinical stage at enrolment are the main predictors of mortality. A similar study was conducted by Zhang *et al.*[4] in Liangshan Prefecture, South-west of China using the same methods. Zhang *et al.*[4] observed that male patients on ART were at a higher risk of death from AIDS than their female counterparts and that a CD4 baseline cell count below 350 cells/mm[3] results in a higher risk of death for the patients than those with a CD4 cell count of at least 350.

Seyoum *et al.*[5] carried out a retrospective cohort study with collected data from clinical records of adult HIV patients following (2006-2010) antiretroviral therapy (ART) in South-western Ethiopia. Their findings reveal the main factors associated with mortality as baseline age (below 35 years) and low adherence to ART.

This study is carried out to explore the survival of HIV patients on combination antiretroviral therapy (cART) using retrospective data from a Wellness clinic in the northern part of South Africa. A multi-state HIV, AIDS, Death (HAD) model is developed and used to compare survival and mortality rates of patients receiving cART. Factors associated with survival or mortality of patients are analyzed using the continuous-time Markov modelling approach. Some of the factors explored include age at baseline, CD4 cell count at baseline, gender, peripheral neuropathy (PN), non-adherence (NA), and an orthogonal viral load residual covariate (RV).

In the next section, the methods used for the analysis of data are explained. Descriptive statistics for the data that is used for analysis are given. This is followed by a presentation of results from the analysis. The last section discusses and concludes the findings.

MATERIALS AND METHODS

Ethical considerations: The procedures used in this study were as approved by the Research Ethics Committee of the University of Venda, South Africa (Protocol number SMNS/13/MBY/01/0625), following the 1964 Helsinki declaration and its subsequent amendments. Additionally, permission to access health facilities was obtained from the Limpopo Provincial Department of Health, South Africa and the collaborating health facilities. Informed consent was obtained from study participants before their involvement and data obtained were stripped of personal identifiers to ensure the anonymity and confidentiality of the participants.

Data: A retrospective cohort study was carried out at a HIV Wellness Clinic Wellness clinic in the Limpopo Province of South Africa on 320 patients living with HIV/AIDS who had been attending ART follow-up care from 2005 to 2009. At treatment commencement (time t = 0), there were 224 females and 96 males. About 50 and 65% of the female and male deaths, respectively occurred during the first 6 months of treatment uptake. The interquartile range of patient ages is (33, 48) years with a mean and median age of 40.62 and 41 years, respectively. The ages were negatively skewed (skew = -0.08) confirming that there were more younger patients than older patients in the cohort.

Time-homogeneous Markov jump model: A Markov jump model on a finite or countable set, S, is a family of random variables $(X_c(t))_{t \geq 0}$ ((right continuous), on a probability space $(\Omega, F_{X_c(s)}, P)$. $F_{X_c(s)}$, denote all the information pertaining to the history of X_c up to s<t and c represents the number of states. According to the continuous-time homogeneous Markov jump process assumptions, individuals transition independently among states. This means we can assume that the transition intensities are constant over time, that is, the transition intensities are independent of t[6,7]. Thus, for the time-homogeneous Markov jump model we have:

$$Q(t) = Q, \forall t$$

where, Q is a c×c transition rate matrix and becomes:

$$q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1c} \\ q_{21} & q_{22} & \cdots & q_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ q_{c1} & q_{c2} & \cdots & q_{cc} \end{pmatrix}$$

For some constant matrix Q. This implies that sojourn time with a particular state, i, has an exponential distribution with rate parameter:

$$-q_{ii} = \lambda_i = \Sigma_{i \neq j} \; q_{ij}$$

where, q_{ij} is the (i, j)th entry of Q. Thus, transition probabilities only depend on the interval between times t_1 and t_2 and not on t_1 itself.

For a continuous-time homogeneous model, the transition probabilities satisfy the Chapman-Kolmogorov's equation[8]:

$$\forall s, t \geq 0, \forall i, j \in S : P_{ij}(t+s) = \sum_{l \in S} P_{il}(t) P_{lj}(s)$$

According to Longini and Hudgens[9]:

$$P\left[X(t+\Delta t) = j | X(t) = 1\right] = \begin{cases} q_{1j}\Delta t + o(\Delta t); 1 \neq j \\ 1 - \lambda_1 \Delta t + o(\Delta t); 1 = j \\ o(\Delta t); \text{otherwise} \end{cases}$$

leads to the Kolmogorov's forward equation:

$$P'_{ij}(t) = \sum_{l \in S} P_{il}(t) q_{lj} = P(t) \, Q, \text{ for all } i, j \in S$$

The Kolmogorov's forward differential equation is derived as:

$$P_{ij}(t+\Delta t) = \sum_{l \in S} P_{il}(t) P_{lj}(\Delta t) \left(\text{Chapman} - \text{Kolmorov's equations}\right)$$

$$= P_{ij}(t) P_{jj}(t) + \sum_{l \neq j} P_{il}(t) P_{lj}(\Delta t)$$

$$= P_{ij}(t) \left(1 + q_{jj}\Delta t + o(\Delta t)\right) + \sum_{l \neq j} P_{il}(t) \left(q_{lj}\Delta t + o(\Delta t)\right)$$

$$= \sum_{l \neq j} P_{il}(t) q_{lj}\Delta t + o(\Delta t) + P_{ij}(t) q_{jj}\Delta t + P_{ij}(t) + o(\Delta t)$$

$$= P_{ij}(t) + \sum_{l \neq j} P_{il}(t) q_{lj}\Delta t + o(\Delta t)$$

Rearranging to derive the forward differential equation gives:

$$\frac{P_{ij}(t+\Delta t) - P_{ij}(t)}{\Delta t} = \sum_{l \neq j} P_{il}(t) q_{lj} + \frac{o(\Delta t)}{\Delta t}$$

Taking the limits as $\Delta t \rightarrow 0$ gives the desired result:

$$P'_{ij}(t) = \sum_{l \neq j} P_{il}(t) q_{lj} \left(\text{since} \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0\right)$$

This is the Kolmogorov forward equation for the process. In the biology literature, this system of equations is termed the chemical master equation[10].

The Kolmogorov backward equation is derived from Chapman-Kolmogorov's equation by substituting s = Δt as follows:

$$P_{ij}(\Delta t + t) = \sum_{l \in S} P_{il}(\Delta t) P_{lj}(t)$$

Now, since:

$$P_{il}(\Delta t) = q_{il}\Delta t + o(\Delta t), \text{ for } l \neq i$$

and:

$$P_{il}(\Delta t) = 1 - \sum_{l \in S} P_{il}(\Delta t) = 1 - \sum_{l \in S} q_{lj}\Delta t + o(\Delta t)$$
$$= 1 + q_{ii}\Delta t + o(\Delta t)$$

We have:

$$P_{ij}(\Delta t + t) = \sum_{l \in S} P_{il}(\Delta t) P_{lj}(t)$$

$$= \sum_{l \neq j} P_{il}(\Delta t) P_{lj}(t) + P_{ii}(\Delta t) \, P_{ij}(t)$$

$$P_{ij}(\Delta t + t) = \sum_{l \neq j} \Delta t q_{il} P_{ij}(t) + \left(1 + \Delta t q_{ii}\right) P_{ij}(t) + o(\Delta t)$$

$$= P_{ij}(t) + \sum_{l \in S} q_{il} P_{lj}(t) \Delta t + o(\Delta t)$$

If we then take $P_{ij}(t)$ term to the left-hand side, divide by Δt and then taking limits.

As Δt→0, we obtain the differential equation:

$$\frac{P_{ij}(\Delta t + t) - P_{ij}(t)}{\Delta t} = \sum_{l \neq j} q_{il} P_{lj}(t) + \frac{o(\Delta t)}{\Delta t}$$

$$P'_{ij}(t) = \sum_{l \neq j} q_{il} P_{lj}(t), \text{ for all } i, j \in S$$

Model formulation: Consider the following HIV, AIDS, Death (HAD) model with the given transition rates. The state space is S = {H, A, D}. These states are based on CD4 cell counts as follows:

$$CD4 = \begin{cases} \text{H; CD4} \geq 200 \text{ (HIV state)} \\ \text{A; CD4} < 200 \text{ (AIDS state)} \\ \text{D; DEATH state} \end{cases}$$
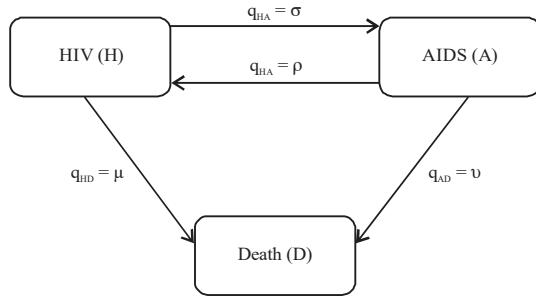
Fig. 1: HIV, AIDS and Death (HAD) model for HIV-infected patients on combination antiretroviral therapy

The CD4 state H represents CD4 cell counts above 200 cells mm^{-3}, state A represents the AIDS-defining state, which is characterised by CD4 cell counts below 200 cells mm^{-3} and state D is the DEATH state that can be reached from either state H or state A. At each stage, an individual is expected to be in either state A, state H or state D. The states H, A and D are mutually exclusive. States H, A and D are defined for patients receiving antiretroviral therapy, such that the transitions between states are bi-directional due to adherence or non-adherence to treatment as shown in Fig. 1.

The transition rate from the AIDS state to the DEATH state is denoted by v. Life may be in the HIV state or the AIDS state on several separate occasions before making the one-way transition to the death state. Alternatively, life may pass from the HIV state to the DEATH state without ever having been in the AIDS state.

Using the notations of the HAD model, the expression for q_{AH}, q_{HH}, q_{HA}, q_{HD}, q_{AD}, q_{AA} and q_{DD} can be given as follows:

$$q_{AH} = \rho, q_{HH} = -(\sigma+\mu) = -\lambda_H \ q_{DD} = \lambda_D = 0,$$
$$q_{HA} = \sigma, q_{HD} = \mu, q_{AD} = v, q_{AA} = -(\rho+v) = -\lambda_A$$

The generator matrix for the HAD model is:

$$Q = \begin{pmatrix} q_{HH} & q_{HA} & q_{HD} \\ q_{AH} & q_{AA} & q_{AD} \\ q_{DH} & q_{DA} & q_{DD} \end{pmatrix} = \begin{pmatrix} -\sigma-\mu & \sigma & \mu \\ \rho & -\rho-v & v \\ 0 & 0 & 0 \end{pmatrix}$$

where, the order of the states has been taken to be H, A then D.

Kolmogorov's forward equation: For the HAD model, the differential equation for P_{HH} (t) can be defined by the general forward equation as a template. This gives:

$$P'_{HH}(t) = \sum_{L \in S} P_{HL}(t) q_{LH}, \text{ for } L = H, A, D$$

$$= P_{HH}(t) q_{HH} + P_{HA}(t) q_{AH} + P_{HD}(t) q_{DH}$$

Now substituting it for transition rates, we have:

$$P'_{HH}(t) = -P_{HH}(t)(\sigma+\mu) + P_{HA}(t)\rho = -P_{HH}(t)\lambda_H + P_{HA}(t)\rho$$

The Kolmogorov's forward equation for the transition probability P_{HA} (t) is:

$$P'_{HA}(t) = \sum_{L \in S} P_{HL}(t) q_{LA}, \text{ for } L = H, A, D$$

$$= P_{HH}(t) q_{HA} + P_{HA}(t) q_{AA} + P_{HD}(t) q_{DA}$$

$$= P_{HH}(t) \sigma - P_{HA}(t)(\rho+v)$$

$$= P_{HH}(t) \sigma - P_{HA}(t) \lambda_A$$

Kolmogorov's backward equation: For the HAD model, the Kolmogorov's backward differential equation for P_{HH} (t) can be obtained using the general backward equation as a template. This gives:

$$P'_{HH}(t) = \sum_{L \in S} q_{HL} P_{LH}(t), \text{ for } L = H, A, D$$

$$= q_{HH} P_{HH}(t) + q_{HA} P_{AH}(t) + q_{HD} P_{DH}(t)$$

Now substituting in the transition rates, we have:

$$P'_{HH}(t) = -(\sigma+\mu) P_{HH}(t) + \sigma P_{AH}(t) = -\lambda_H P_{HA}(t) + \sigma P_{AH}(t)$$

The backward equation for the transition probability P_{HA} (t) is given by:

$$P'_{HA}(t) = q_{HH} P_{HA}(t) + q_{HA} P_{AA}(t) + q_{HD} P_{DA}(t)$$

$$P'_{HA}(t) = -(\sigma+\mu)P_{HA}(t) + \sigma P_{AA}(t) = -\lambda_H P_{HA}(t) + \sigma P_{AA}(t)$$

Maximum likelihood estimators: From the HAD model defined:
Let:

T_{hi} = Waiting time of the ith life in the HIV state
T_{ai} = Waiting time of the ith life in the AIDS state
S_i = Number of transitions HIV→AIDS by the ith life
R_i = Number of transitions AIDS→HIV by the ith life
D_i = Number of transitions HIV→Death by the ith life
U_i = Number of transitions AIDS→Death by the ith life

We also need to define totals:

$$T_H = \sum_{i=1}^{N} T_{Hi}, \ T_A = \sum_{i=1}^{N} T_{Ai}, \ S = \sum_{i=1}^{N} S_i, \ R = \sum_{i=1}^{N} R_i, \ D = \sum_{i=1}^{N} D_i, \ U = \sum_{i=1}^{N} U_i$$

Using the lower case symbols for the observed samples, it can be shown that the likelihood for the parameters, μ, ν, σ, ρ for the HAD model is given by:

$$L(\mu, \nu, \sigma, \rho) = \prod_{i=1}^{N} e^{-(\mu+\sigma)T_{Hi}} e^{-(\nu+\rho)T_{Ai}} \mu^{d_i} \nu^{u_i} \sigma^{s_i} \rho^{r_i}$$
$$= e^{-(\mu+\sigma)\sum T_{Hi}} e^{-(\nu+\rho)\sum T_{Ai}} \mu^{\sum d_i} \nu^{\sum u_i} \sigma^{\sum s_i} \rho^{\sum r_i}$$
$$= e^{-(\mu+\sigma)T_H} e^{-(\nu+\rho)T_A} \mu^d \nu^u \sigma^s \rho^r$$

Where:

$$d = \sum d_i, \ u = \sum u_i, \ s = \sum s_i, \ r = \sum r_i$$

The likelihood function L (μ, ν, σ, ρ) for the ith life reflects:

- Probability of the life remaining in the HIV state for total time T_{Hi} and in the AIDS state for time T_{Ai}, giving the factors $e^{-(\mu+\sigma)T_{Hi}}$ and $e^{-(\nu+\rho)T_{Ai}}$ respectively
- Probability of life making the relevant number of transitions between states giving the factors $\mu^{d_i}, \nu^{u_i}, \sigma^{s_i}$ and ρ^{r_i}

The likelihood factorises into functions of each parameter of the form $e^{-\mu T_A} \mu^d$:

$$L(\mu, \nu, \sigma, \rho) = e^{-(\mu+\sigma)T_H} e^{-(\nu+\rho)T_A} \mu^d \nu^u \sigma^s \rho^r$$
$$= \left(e^{-\mu T_H} \mu^d\right) \times \left(e^{-\sigma T_H} \sigma^s\right) \times \left(e^{-\nu T_A} \nu^u\right) \times \left(e^{-\rho T_A} \rho^r\right)$$

So the log-likelihood is:

$$\log L = -(\mu+\sigma)T_H - (\nu+\rho)T_A + d \log \mu + u \log \nu + s \log \sigma + r \log \rho$$

Differentiating with respect to each of the four parameters gives:

$$\frac{\partial \log L}{\partial \mu} = -T_H + \frac{d}{\mu}, \ \frac{\partial \log L}{\partial \nu} = -T_A + \frac{u}{\nu}$$

$$\frac{\partial \log L}{\partial \sigma} = -T_H + \frac{s}{\sigma}, \ \frac{\partial \log L}{\partial \rho} = -T_H + \frac{r}{\rho}$$

Setting each of the derivatives to 0 and solving the resulting equations, we see that:

$$\hat{\mu} = \frac{d}{T_H}, \ \hat{\nu} = \frac{u}{T_A}, \ \hat{\sigma} = \frac{s}{T_H}, \ \hat{\rho} = \frac{r}{T_A}$$

When there is more than one parameter to be estimated, the second-order condition to check for maxima is that the Hessian matrix is negative definite, or equivalently, the

eigenvalues of the Hessian matrix are all negative. The Hessian matrix is the matrix of the second derivatives. So, in this case, we consider the matrix:

$$\begin{pmatrix} -\dfrac{d}{\mu^2} & 0 & 0 & 0 \\ 0 & -\dfrac{u}{\nu^2} & 0 & 0 \\ 0 & 0 & -\dfrac{s}{\sigma^2} & 0 \\ 0 & 0 & 0 & -\dfrac{r}{\rho^2} \end{pmatrix}$$

Since this is a negative definite matrix, the maximum likelihood estimates of μ, ν, σ, ρ are:

$$\hat{\mu} = \frac{d}{T_H}, \ \hat{\nu} = \frac{u}{T_A}, \ \hat{\sigma} = \frac{s}{T_H}, \ \hat{\rho} = \frac{r}{T_A}$$

Therefore, the maximum likelihood estimators are as given:

$$\hat{\mu} = \frac{D}{T_H}, \ \hat{\nu} = \frac{U}{T_A}, \ \hat{\sigma} = \frac{S}{T_H}, \ \hat{\rho} = \frac{R}{T_A}$$

Coding of covariates: The effects of covariates on estimated transition intensities are analysed. This helps in determining variables that have a strong influence on the survival of HIV patients receiving antiretroviral therapy. These variables include age, peripheral neuropathy (PN), non-adherence (NA), gender, an orthogonal viral load residual covariate (RV). As proposed by Shoko *et al*.[11] the effect of the orthogonal viral load variable is included in the continuous-time Markov model. The variables are coded as follows:

$$Age = \begin{cases} 1, \leq 45 \text{ years} \\ 0, > 45 \text{ years} \end{cases}, PN = \begin{cases} 1, \text{Yes} \\ 0, \text{No} \end{cases}, NA = \begin{cases} 1, \text{Yes} \\ 0, \text{No} \end{cases}, Gender = \begin{cases} 1, \text{male} \\ 0, \text{female} \end{cases},$$

$$\text{Orthogonal viral load residual } (RV) = \begin{cases} 1, \text{if positive} \\ 0, \text{if negative} \end{cases}$$

The residuals are obtained from a linear regression model of viral load on CD4 cell count. The residual covariate is independent of the CD4 count covariate correcting for collinearity effects. For more details on the collinearity argument, see Shoko *et al*.[11]. The effects of the covariates on transition intensities, q_{ij}, for a patient h is given by the model:

$$q_{ij} = \begin{array}{c} q_{ij}^{(0)} \exp (\beta_{ij}^{(Age)} Age_h + \beta_{ij}^{(PN)} PN_h \\ + \\ \beta_{ij}^{(Gender)} Gender_h + \beta_{ij}^{(NA)} NA_h + \beta_{ij}^{(RV)} RV_h) \end{array}$$

For this model, the baseline transition intensities, $q_{ij}^{(0)}$, refer to a patient with age category 0 (over 45 years old), no LA, no PN, Gender = 0 (female), no NA and RV = 0 means lower viral load than expected. The transition intensities, q_{ij}, are presented in rates per year. q_{ij} are the elements of a 3×3 transition intensity matrix Q from a continuous time-homogeneous Markov process.

RESULTS

This section gives the results from the analysis of the data based on the HAD model given earlier in Fig. 1 of the methods section. Three states are defined namely: state H-HIV state marked by CD4 cell count greater than 200 cell mm^{-3}, state A-the AIDS-defining state, marked by CD4 cell count below 200 cells/mm^3 and state D-Death. We start off by computing the transition counts. The transition counts are shown in Table 1.

Results from Table 1 above show fewer deaths recorded from CD4 cell count greater or equal to 200 (3.1%) than from the AIDS-defining state (15.6%). Once a patient has achieved a CD4 cell count greater or equal 200 (state H), there is a possibility of reverting to the AIDS state.

Next, the maximum likelihood estimators for the HAD model are computed using the continuous-time homogeneous Markov model.

Continuous-time homogeneous Markov model for the HAD model: In this subsection, we give estimates of the transition intensity matrix given by Q_{ij}, probability of each state being next, given by:

$$p_{ij} = \frac{q_{ij}}{\sum_{i \neq j} q_{ij}} = \frac{q_{ij}}{\lambda_i} \text{ for } i \neq j$$

For example, in Table 2:

$$p_{HA} = \frac{q_{HA}}{q_{HA} + q_{HD}} = \frac{0.2273}{0.2273 + 0.0201} = 0.9188$$

The results are shown in Table 2.

Results from Table 2 show higher risks of death from the AIDS state (state A) than from state H where the CD4 cell count is above 200. The results also showed that for patients on antiretroviral therapy and in the AIDS-defining state, there are higher rates of recovery to a state of CD4 cell count above 200 cell mm^{-3} compared to rates of transitions to the death state. Thus, confirmation that antiretroviral therapy improves the survival of HIV infected patients.

Next, we compute the total time spent in each state and the mean sojourn time for the HAD model as shown in Table 3.

The total times spent in each state reveal higher survival chances once the HIV-infected patient is alive and has transitioned out of the AIDS-defining state. Results from Table 3 show that patients with CD4 cell count above 200 and on cART can spend approximately 17 years before absorption into the death state. HIV-infected patients can also spend an average of about 4 years in state H before the transition to other states.

For the continuous-time homogeneous model with parameters estimated above, we plot the percentage prevalence in each of the states indicated in the HAD model for HIV-infected patients on combination antiretroviral therapy.

The prevalence plots from Fig. 2 show that the fitted model overestimates observed percentage prevalence in state 1 (state H). The fitted model gives a better fit of the observed patients in the AIDS-defining state 2 (state A). Observed death prevalence is underestimated by the fitted model. This suggests that transition intensities may not be constant throughout the whole study period. Hence, the need to fit a continuous-time nonhomogeneous Markov model. This is done below.

Table 1: Transition counts for the HAD model

From	To		
	H	A	D
H	721	67	10
A	230	851	50
D	0	0	0

Table 2: Maximum likelihood estimates of the transition intensities, probability of each state being next and transition probabilities for the HAD model (95% CI are given in brackets)

Transitions	q_{ij}	p_{ij}	$P_{ij}(t)$
State H to H	-0.2474 (-0.3104, -0.1972)	0	0.8292
State H to A	0.2273 (0.1781, 0.2902)	0.9188 (0.7978, 0.9703)	0.1420
State A to H	0.6364 (0.5577, 0.7261)	0.8319 (0.7851, 0.8721)	0.3973
State A to A	-0.7650 (-0.8619, -0.6789)	0	0.5062
State H to D	0.0201 (0.0075, 0.0541)	0.0812 (0.0297, 0.2022)	0.0288
State A to D	0.1286 (0.0957, 0.1727)	0.1681 (0.1279, 0.2149)	0.0965

-2* log-likelihood = 1977.895

Table 3: Estimates of the total time spent in each state and sojourn times (95% CI are given in brackets)

Transitions	Total time spent (h) in each state	Sojourn time (h)
State H	17.1557	4.0416 (3.2211, 5.0711)
State A	5.0986	1.3073 (1.1602, 1.4730)
State D	Infinity	

Table 4: Maximum likelihood estimates for the baseline transition intensities and hazard ratios for each time interval (95% CI are given in brackets)

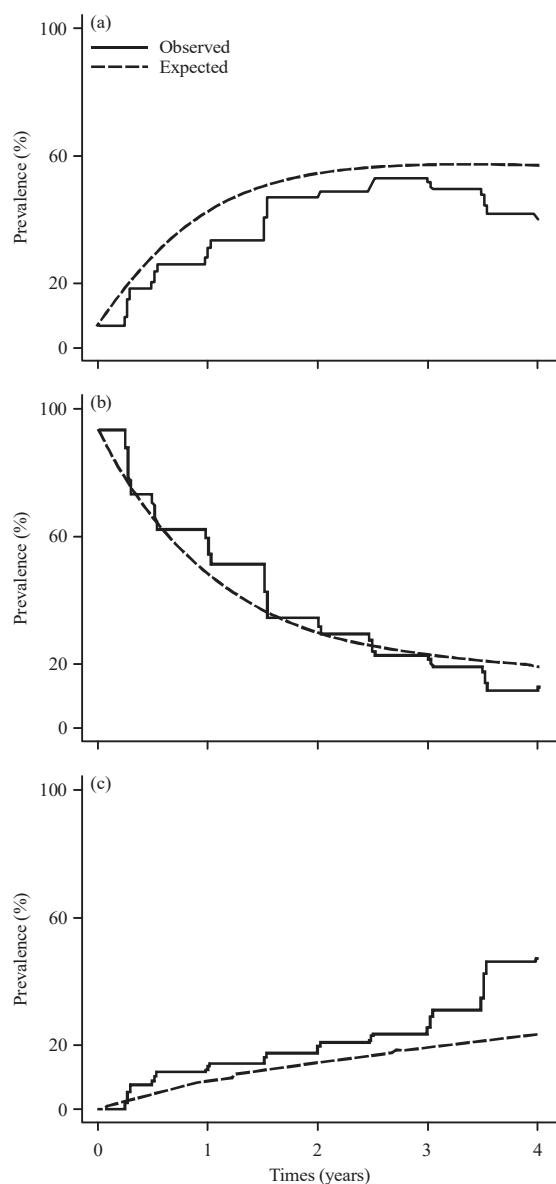| Transitions	$q_{ij}^{(0)}$	HR (0.5, 1 year)	HR (1, Infinity years)
State H to A	0.2722 (0.2049, 0.3616)	0.6142 (0.2336, 1.6145)	0.4359 (0.2046, 0.9286)
State A to H	0.6725 (0.5856, 0.7721)	0.6994 (0.4538, 1.0779)	1.0304 (0.7486, 1.4184)
State H to D	0.0470 (0.0207, 0.1070)	0.3199 (0.0347, 2.9479)	0.0730 (0.0102, 0.5207)
State A to D	0.0892 (0.0569, 0.1396)	0.1836 (0.0382, 0.8832)	0.2750 (0.1186, 0.6373)

-2* log-likelihood = 1935.603

Fig. 2(a-c): Prevalence for the continuous-time homogeneous Markov plots for the HAD model (a) State 1, (b) State 2 and (c) State 3

Next, we fit a continuous-time inhomogeneous Markov model to assess the interval in which immune deterioration is experienced after patients have achieved an improvement in the immune system.

Results from Table 4 showed the maximum likelihood estimates for the baseline transition intensities and hazard ratios for the effects of time on transition intensities. The results show that in the second 0.5 years of treatment uptake, there is a reduction in the rates of immune suppression. The second 0.5 years of treatment uptake is also characterised by a reduction in deaths particularly from the AIDS-defining state. From 1 year of treatment uptake onwards, there is a significant increase in the rates of immune recovery from the AIDS-defining state. However, there is a slight increase in the rate of occurrences of deaths after 1 year of treatment compared to the 0.5 years to 1-year interval. There is a further reduction in deaths for patients with CD4 cell count above 200, 1-year post-treatment uptake compared to the 0.5 years to 1-year interval. This means that chances of survival for HIV-infected patients increase once their CD4 cell count is above 200 and the effect of time for this to happen is crucial.

We now plot percentage prevalence for each of the states in the continuous-time non-homogeneous model. The plots are shown in Fig. 3.

Results in Fig. 3 show a great improvement in fitting a non-homogeneous model compared to the homogeneous model. The fitted model in Fig. 3 now gives an almost perfect fit of the observed percentage prevalence. In the next section, we analyse factors that contribute to the survival of HIV-infected patients on combination antiretroviral therapy.

Covariates that contribute to the survival of HIV-infected patients on cART: In this subsection, we analyse the effects of routinely collected viral load (an orthogonal viral load covariate (RV)), age, gender, peripheral neuropathy (PN) and non-adherence to treatment (NA) on the survival of HIV-infected patients on cART.

Results from Table 5 show that risks of immune suppression from CD4 cell count above 200 to the AIDS-defining state is attributed to non-adherence to treatment and having higher viral load than expected. Patients below the age of 40 years had significantly higher rates of immune recovery than their older counterparts. Results also show that male patients, patients below the age of 40 years, patients with a higher viral load than expected and non-adherent patients are at higher risk of death from the CD4 cell count level above 200.
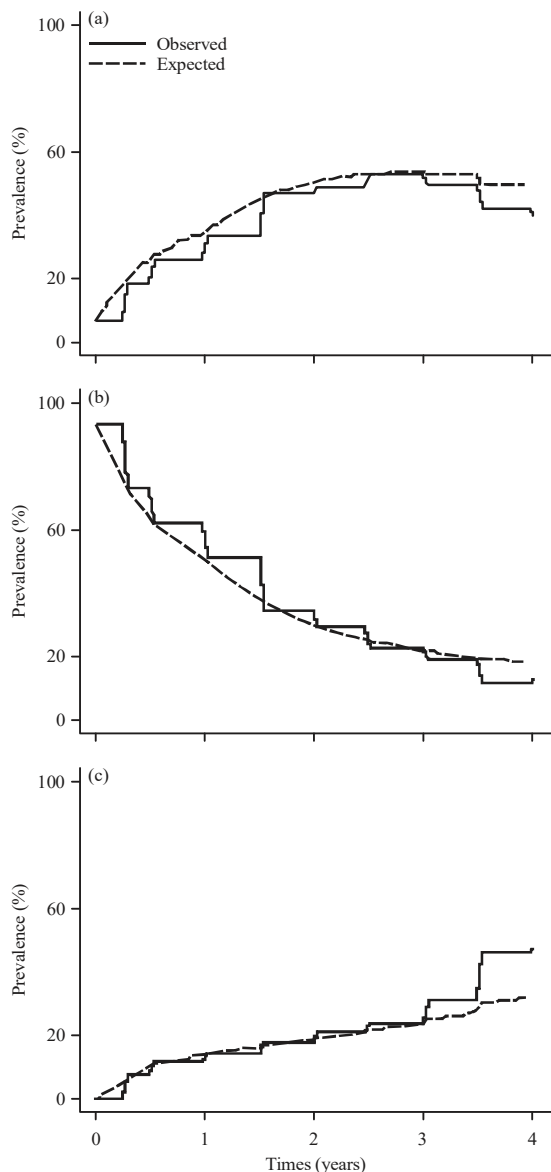
Fig. 3(a-c): Prevalence (%) for the continuous-time non-homogeneous Markov model for the HAD model (a) State 1, (b) State 2 and (c) State 3
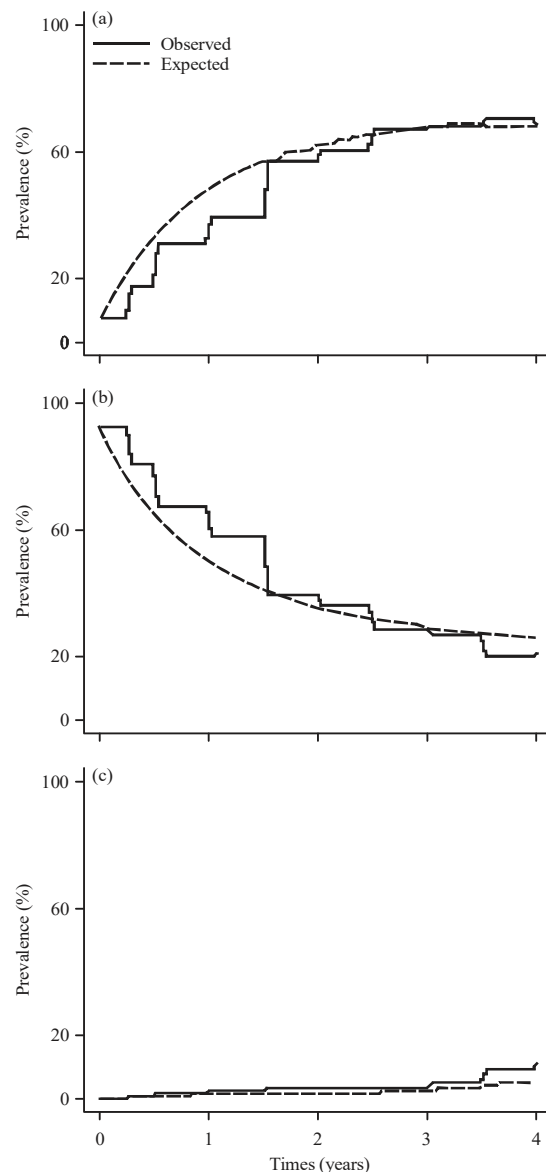

Fig. 4(a-c): Prevalence (%) for the continuous-time-homogeneous Markov model with covariates (a) State 1, (b) State 2 and (c) State 3

Figure 4 shows that the inclusion of the effects of covariates in the Markov model results in a good fit of the observed percentage prevalence for each of the states in the HAD model. This has also resulted in better prediction of mortality (state 3). This is shown by death prevalence that accumulated to slightly below 20% by the end of the 4 years of treatment uptake. This gives the best prediction of mortality since 60 deaths out of 320 patients occurred during the study. This is approximately equal to 18.8%.

Assessment of the fitted models: In this section, we assess the fitted models by performing a likelihood ratio test as well as comparing the -2 *log-likelihood(-2LL) ratios for the fitted models. The results are given in Table 6.

Results from Table 6 show that the likelihood ratio test for the time-homogeneous model versus the time-homogeneous model with covariates is in favour of a time-homogeneous model with covariates. A comparison of the nonhomogeneous model with the homogeneous model with covariates is in favour of the homogeneous model with covariates.

Table 5: Maximum likelihood estimates for the baselines transition intensities with hazard ratios for each covariate (95% CI are given in brackets)

Transitions	Baseline	RV	Age	Gender	PN	NA
State H to A	0.1779 (0.1242, 0.2547)	11.147 (2.6110, 47.588)	0.8090 (0.4539, 1.442)	0.8162 (0.4516, 1.4749)	0.8334 (0.4101, 1.694)	2.2005 (1.050, 4.612)
State A to H	0.7076 (0.6143, 0.8152)	0.928 (0.6466, 1.332)	1.1476 (0.8383, 1.571)	0.6648 (0.4831, 0.9148)	0.6826 (0.4645, 1.003)	0.59583 (0.36840, 0.9638)
State H to D	0.0024 (0.0002, 0.2287)	1.091 (0.0914, 13.026)	34.562 (0.00001, 1756)	1.7609 (0.1488, 20.8415)	0.0650 (0.00005, 84142)	3.80314 (0.3184, 45.43)
State A to D	0.0012 (0.00007, 0.1909)	0.380 (0.0599, 2.411)	0.1001 (0.0062, 1.605)	0.0071 (0.00001, 278)	0.0145 (0.0002, 8519)	0.0319 (0.00005, 2119)

-2 * log-likelihood = 1391.64

Table 6: Likelihood ratio tests for the fitted models

Models tested in pairs	Preferred model	-2 log LR	Degree of freedom	p-value
Homogeneous with no covariates and homogeneous with covariates)	Homogeneous with covariates	586.2517	20	<0.0001
Non-homogeneous and homogeneous with no covariates	Nonhomogeneous model	42.29228	8	<0.0001
Non-homogeneous and homogeneous with covariates	Homogeneous with covariates	543.9594	12	<0.0001

Time-homogeneous model: -2LL = 1977.895, Time-nonhomogeneous model: -2LL = 1935.603, time-homogeneous model with covariates: -2LL = 1391.64

The time-homogeneous model with covariates also has the lowest-2*log-likelihood compared to all the other fitted models. This shows that the survival of patients on cART is best explained by a time homogenous model with covariates including the routinely collected viral load covariate.

DISCUSSION

In this study, the survival of HIV-infected patients after ART initiation is analysed. Continuous-time Markov models are fitted based on the states from an HIV, AIDS and DEATH (HAD) states model. These states are based on CD4 cell count. Factors that affect the survival of HIV-infected patients on ART were analysed. These, among others, include age, gender, routinely collected viral load, non-adherence, peripheral neuropathy and time on cART.

Results from Table 5 on estimated hazard ratios showed that patients below the age of 40 years achieved a normal CD4 cell count 1.15 times faster than their older counterparts. This is corroborated by findings from a previous study which concluded lower mean CD4 increases for older patients than younger patients[12]. However, these patients have higher risks (Hazard ratio = 34.6, CI: (0.00001, 1756)) of mortality compared to their older counterparts.

Time on treatment also had a significant effect on the survival of patients. Thus, after one year of treatment uptake, patients have accelerated transitions from the AIDS state to the HIV state defined by CD4 cell count above 200 cell/mm³.

Males have generally lower chances of survival than their female counterparts. Results showed that males also have slower rates of immune recovery than females. Sieleunou *et al.*[3] also concluded that gender is a predictor of mortality for patients on antiretroviral therapy.

Non-adherence to treatment reduces the chances of survival for patients on cART. Non-adherence accelerates (Hazard ratio = 2.2) transitions from CD4 cell count state above 200 cell/mm³ to the AIDS state defined by a CD4 cell count below 200 cells/mm³. Patients who were non-adherent

to treatment are 3.8 times more likely to transit from the CD4 state above 200 cell/mm³ to death compared to patients who were adherent to treatment. This is corroborated by the findings from Seyoum *et al.*[5] who observed that low-adherence to ART is associated with increased mortality.

Patients with higher viral loads than expected had higher risks (Hazard ratios = 11.147, CI: (2.6110, 47.588)) of HIV progression to the AIDS state and high risks (Hazard ratio = 1.091, CI: (0.0914, 13.026)) of mortality from a CD4 cell count above 200 cell/mm³. Thus, due to the effects of treatment and time spent on cART, patients are expected to have suppressed viral loads (possibly undetectable viral load) leading to an improved immune system. However, if the viral load remains higher than expected, this increases risks of immune deterioration even after achieving normal CD4 cell counts and consequently, mortality risks are increased.

A continuous-time homogeneous model without covariates, a continuous-time nonhomogeneous model and a continuous-time homogeneous model with covariates were fitted for the data. These models were assessed to select the model that best predicts the survival of HIV-infected patients on cART. Form both the continuous-time homogeneous model and the nonhomogeneous model, cumulative deaths by the end of four years of treatment uptake had increased to close to 40%. For the model with covariates, cumulative deaths prevalence had increased to below 20%. However, from the data only 60 out of 320 deaths were recorded throughout the whole study period which is approximately equal to 18.8%. Therefore, we conclude that a continuous-time homogeneous Markov model with the effects of covariates, including the orthogonal viral load covariate, gives the best prediction of survival of HIV-infected patients on cART than the other two models. This corroborates with the findings from studies by Shoko *et al.*[11] who proposed the inclusion of both viral load monitoring and CD4 cell count in one model to account for the aspect of mortality which one variable can fail to account for without the inclusion of the other.

CONCLUSION

Recovery from AIDS state by HIV infected patients on cART is likely to occur after one year of cART treatment. However, if the viral load remains higher than expected, this increases risks of immune deterioration even after having achieved normal CD4 cell counts and consequently, mortality risks are increased.

SIGNIFICANCE STATEMENT

This study concludes that constant monitoring of an HIV/AIDS patient's viral load can be beneficial for reducing the risks of immune deterioration and mortality. This study will help the researchers to uncover the critical areas regarding HIV/AIDS progression and causes of mortality that many researchers were not able to explore. Thus, a new theory on HIV, AIDS, Death (HAD) Markov model may be arrived at.

REFERENCES

1. Hogg, R.S., B. Yip, K.J. Chan, E. Wood, K.J. Craib, M.V. O'Shaughnessy and J.S. Montaner, 2001. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. J. Am. Med. Assoc., 286: 2568-2577.
2. Moshago, T., D.B. Haile and F. Enqusilasie, 2014. Survival analysis of HIV infected people on antiretroviral therapy at Mizan-Aman General Hospital, Southwest Ethiopia. Int. J. Sci. Res., 3: 1462-1469.
3. Sieleunou, I., M. Souleymanou, A.M. Schönenberger, J. Menten and M. Boelaert, 2009. Determinants of survival in AIDS patients on antiretroviral therapy in a rural centre in the Far North Province, Cameroon. Trop. Med. Int. Health, 14: 36-43.
4. Zhang, G., Y. Gong, Q. Wang, L. Deng and S. Zhang *et al.*, 2016. Outcomes and factors associated with survival of patients with HIV/AIDS initiating antiretroviral treatment in Liangshan Prefecture, Southwest of China: A retrospective cohort study from 2005 to 2013. Medicine, Vol. 95, No. 27. 10.1097/MD.0000000000003969.
5. Seyoum, D., J.M. Degryse, Y. Kifle, A. Taye and M. Tadesse *et al.*, 2017. Risk factors for mortality among adult HIV/AIDS patients following antiretroviral therapy in Southwestern Ethiopia: An assessment through survival models. Int. J. Environ. Res. Public Health, Vol. 14, No. 3. 10.3390/ijerph14030296.
6. Chiang, C.L., 1968. Introduction to Stochastic Processes in Biostatistics. Wiley, New York.
7. Kay, R., 1986. A Markov model for analysing cancer markers and disease states in survival studies. Biometrics, 42: 855-865.
8. Isham, V., 2009. Lecture notes for LTCC course on stochastic processes. LTCC-Stochastic Processes, VSI, 2008-2009. https://pdfs.semanticscholar.org/0a45/39aec979e4574e3f3db3db1bfcbb80518dbf.pdf
9. Longini, I.M.Jr. and M.G. Hudgens, 2003. Lecture notes on stochastic processes in biostatistics: Applications to infectious diseases. https://pdfs.semanticscholar.org/a984/00afd1e90592d35c58065a041bc9ba36a349.pdf
10. Serfozo, R., 2009. Basics of Applied Stochastic Precesses. (Probability and its Application). Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-642-43043-5, Pages: 443.
11. Shoko, C., D. Chikobvu and P.O. Bessong, 2018. A Markov model to estimate mortality due to HIV/AIDS using viral load levels-based states and CD4 cell counts: A principal component analysis approach. Infect. Dis. Ther., 7: 457-471.
12. Maskew, M., A.T. Brennan, D. Westreich, L. McNamara, A.P. MacPhail and M.P. Fox, 2012. Gender differences in mortality and CD4 count response among HIV-positive patients virally suppressed within 6 months of antiretroviral therapy initiation. J. Women's Health, 22: 1-9.