Asian Journal of Probability and Statistics

Volume 21, Issue 4, Page 1-13, 2023; Article no.AJPAS.97022
ISSN: 2582-0230

Estimating Sojourn Time and Transition between Clinical States of HIV Patients under ART Follow up in Namibia

Simon Pombili Kashihalwa [a*], Lilian Pazvakawambwa [b] and Josua Mwanyekange [a]

[a] Department of Mathematics, Statistics and Actuarial Science, Namibia University of Science and Technology, Namibia.
[b] Department of Statistics and Population Studies, University of Namibia, Namibia.

Authors' contributions

All authors contributed equally to the production of the manuscript. All authors have read and agreed to the published version of the manuscript.

Article Information

DOI: 10.9734/AJPAS/2023/v21i4468

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/97022

Original Research Article

Received: 02/01/2023
Accepted: 08/03/2023
Published: 17/03/2023

Abstract

Background: Sojourn time refers to the amount of time a HIV patient spends in each clinical state in a single stay before he/she makes a transition to another state. HIV can be broken down into a number of intermediate states, based on CD4 counts. The four states of the Markov process of HIV are commonly defined as: S1: CD4 count > 500 cells/microlitre of blood; S2: 350 < CD4 count ≤ 500 cells/microlitre of blood; S3: 200 < CD4 count ≤ 350 cells/microlitre of blood; S4: CD4 count ≤ 200 cells/microliter of blood.
Aims: The aim of the study was to estimate sojourn and transition between clinical states of patients under ART in Namibia using homogenous semi-Markov processes, on data obtained from MoHSS.
Methods: A retrospective study design was used to obtain data on 2422 patients who were observed 11028 times, during 2008 to 2017 follow up period. The four staged semi-Markov model was employed to estimate sojourn times and transition between clinical states.

**Corresponding author: Email: skashihalwa@yahoo.com, skashihalwa@nust.na;*

Results: Results indicates that 1637 (67.6%) were female and 785 (32.41%) were male .657(27.13%) patients started ART in state 1, 683(28.19%) patients started ART in state 2, 677(27.95%) patients started ART in state 3 and 405(16.72%) patients started ART in state 4, at treatment commencement ($t = 0$). As expected, the probabilities of transiting from good to worse states increased with time. After 6 months, the probabilities of transiting from state 1 to 3, and from state 1 to 4 are 0.023 and 0.004 respectively. Whereas after 12 months, the probabilities of transiting from state 1 to 3, and from state 1 to 4 are 0.059 and 0.010 respectively. As time increased the probabilities to remain in the same state is decreasing (probabilities of remaining in state 1 after 6, 12 and 18 months is 0.804, 0.698 and 0.633). Sojourn times for states 1, 2, 3 and 4 were 22, 8, 10 and 15 months respectively.

Conclusions: Sojourn time is of interest in HIV modeling, as it gives a signal of how HIV is progressing. Longer sojourn times indicates slow HIV progression and shorter sojourn times indicates rapid HIV progression. As time increases, transition probabilities from good states to worse states increases.

Keywords: Stochastic; semi-Markov processes; multi-state model; CD4; progression.

1 Introduction

HIV disease is one of the leading causes of death in Namibia and worldwide [1]. HIV disease is not only one of the leading causes of disability and human misery, but it also has a significant negative economic impact due to lost productivity and increased healthcare costs. [2]. CD4 cell counts is important in understanding the transition from one clinical state to another. Depreciation of the CD4 cell counts leads to transiting to worst states. The need to estimate the sojourn time and transition between clinical states prompted this study.

The 4 states of the Markov process of HIV illness based on CD4 are commonly defined as: S1: CD4 count > 500 cells/microlitre of blood; S2: 350 < CD4 count ≤ 500 cells/microlitre of blood; S3: 200 < CD4 count ≤ 350 cells/microlitre of blood; S4: CD4 count ≤ 200 cells/microlitre of blood [3]. It is therefore important to understand the natural history and the amount of time a patient spent in each clinical state. As time spent in each state of the disease cannot be estimated based on clinical and immunological measures, this needs to be modeled by the semi-Markov stochastic process [3,4]. A semi-Markov process is defined as a stochastic process, which can be in any state. Each time it enters a state, it remains there for an unpredictable period of time before potentially moving forward or backward into another state [5].

Recent studies on HIV have estimated sojourn times and transition between clinical states, using homogenous semi-Markov processes. Dynamical models were used to model the progression of HIV [6-8]. Homogeneous semi-Markov processes have been used to estimate the proportion of individuals changing from one clinical state to another [9-11]. Semi-Markov models were applied to HIV disease evolution and compared sojourn time distributions, exponential and Weibull probability distributions [12]. Moreover, other authors applied the homogenous Markov process to HIV disease under a combination treatment therapy [13]. Whiles, there is some progress made towards HIV progression, there is still much to be done, due to sojourn time and transition between clinical states of HIV patients hence the import of this study.

The aim of the study was to estimate sojourn and transition between clinical states of patients under ART in Namibia using four staged homogenous semi-Markov processes.

The next section explores the materials and methods of Markov modeling and an illustrative case study on sojourn time and transition between HIV clinical states. In this section, data used in the analysis is described and formulation of the model based on the data is explained. This is followed by a section on the results and discussions. The final section concludes on the findings.

2 Materials and Methods

2.1 Study area, design and data collection

This retrospective cohort study was conducted in Namibia, from January 2008- January 2012 to December 2017. All registered patients who were infected with HIV and whose CD4 counts were measured at least once,

were included in this study. The study involved the observation of 2422 patients, who were seen 11028 times in total. The semi-Markov model was employed to predict transitional probabilities and sojourn times. At treatment commencement ($t = 0$), 657(27.13%) patients started ART in state 1, 683(28.19%) patients started ART in state 2, 677(27.95%) patients started ART in state 3 and 405(16.72%) patients started ART in state 4. Fig. 1 indicates the 4 states an HIV infected patient may go through. The arrows in Fig. 1 represents communication between states. All states are inter-related.

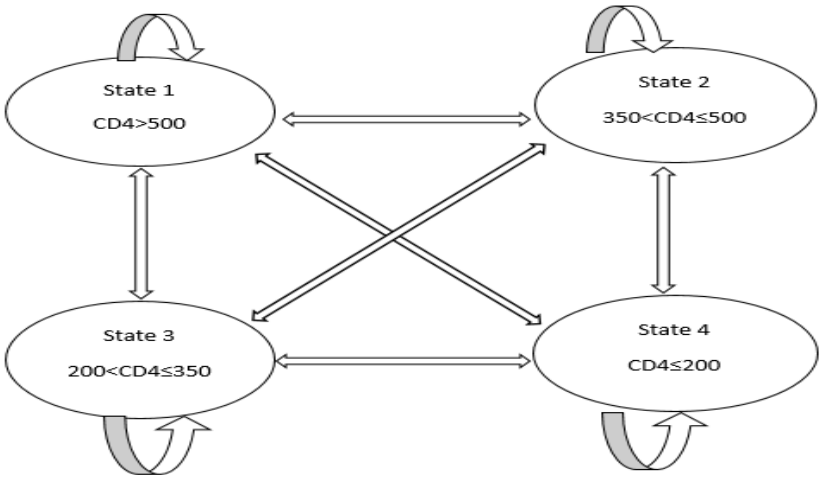

Fig. 1. Transition diagram

2.2 Modelling homogenous semi-Markov processes

Markov chain is described as follows: there exist a set of state, $S = \{s_1, s_2, s_3, s_4\}$. The processes start in one of this state and make a transition from one state to another state. If the chain is in state s_{ij} then it move to state s_j with the probability of p_{ij} or stay in the same state with the probability of p_{ii} [14].

One of the Markov assumptions is that future development only depends on the current state, not on the previous states and the current state should include all relevant history [15]. This assumption imposes restrictions on the distribution of the sojourn time in a state, which should be exponentially distributed, in case of continuous-time Markov process and geometrically distributed, in case of a discrete-time Markov process [16].

To overcome this, the Markov assumption must be relaxed in order to allow exponential distributed sojourn times in any state and still have the Markov assumption, but in a more flexible manner, since this paper deal with continuous semi-Markov the distribution of sojourn time is exponential [17]. The resultant process based on these properties, is known as a semi-Markov process. A semi-Markov process is concerned with the random variables, that describe the state of the process at some time and it is a generalization of the Markov process.

A semi-Markov process is that makes transitions from state to state like a Markov process. However, the amount of time spent in each state before a transition to the next state, occurs is an arbitrary random variable [18]. In this study, a homogenous semi-Markov was adopted for predicting sojourn times and transition matrix using longitudinal CD count measurements.

Homogeneous semi-Markov processes (HSMP) were introduced in the 1950s, independently, with the objective of generalizing Markov processes [19,20]. A homogenous semi-Markov process (HSMP) model is define as follows [13]:

Let $X_n : \Omega \to S$ be a stochastic process with state space $S = \{S_1, S_2, \dots, S_m\}$ and $T_n : \Omega \to \mathbb{R}$ be the time of the n^{th} transition, with Ω domain of the process and $\mathbb{R}$ set of real numbers. At this point the time is a random variable. The kernel $Q = [Q_{ij}]$ associated with the process and the transition probability P_{ij} of the embedded Markov chain are defined as follows:

$$Q_{ij}(t) = P[T_{n+1} = j, T_{n+1} - T_n \leq t | X_n = i] \tag{2.2.1}$$

The probability of moving from state *i* to state *j* is given by

$$P_{ij} = \lim_{t\to\infty} Q_{ij}(t) \tag{2.2.2}$$

Define the probability that the process will leave a state *i* in a time *t* as

$$H_i(t) = P[T_{n+1} - T_n \leq t | X_n = i] = \sum_{j=1}^{m} Q_{ij}(t) \tag{2.2.3}$$

The distribution of waiting time in each state *i*, given that the state *j* is subsequently occupied is

$$G_{ij}(t) = P[T_{n+1} - T_n \leq t | X_n = i, X_{n+1} = j], \tag{2.2.4}$$

which can be computed as:

$$G_{ij}(t) = \begin{cases} \frac{Q_{ij(t)}}{P_{ij}}, & if P_{ij} \neq 0 \\ 1, & if P_{ij} = 0 \end{cases} \tag{2.2.5}$$

For any homogenous semi-Markov process $\{X(t), t \geq 0\}$, the transition probabilities are given by equation (2.3.6), for which the solution should be obtained, using the progression equation (2.3.7).

$$\emptyset_{(ij)}(t) = P[X(t) = j | X(0) = i], \tag{2.2.6}$$

$$\emptyset_{ij}(t) = (1 - H_i(t))\delta_{ij} + \sum_{l=1}^{m} \int_0^t Q_{il}(\tau)\phi_{lj}(t-\tau)\, d\tau \tag{2.2.7}$$

At this point δ_{ij} represents the Kronecker delta δ. An approximate solution of equation (2.3.7) can be obtained using the general numerical integration formula given in [21]. In the same paper, they proved that the numerical solution of the process converges to the discrete time HSMP described as an infinite countable linear system:

$$\phi_{ij}^h(kh) = d_{ij}^h(kh) + \sum_{i=1}^{m} \sum_{\tau=1}^{k} v_{ij}^h(\tau h)\phi_{ij}^h((k-\tau)h) \tag{2.2.8}$$

where *h* stands for the step measure of the approximation and

$$d_{ij}^h(kh) = \begin{cases} 0 & if\ i \neq j, \\ 1 - H_i^h(kh), & if\ i = j, \end{cases} \tag{2.2.9}$$

$$v_{ij}^h(kh) = \begin{cases} 0, & if\ i \neq j \\ \varrho_{ij}^h(kh) - \varrho_{ij}^h\big((k-1)h\big), & if\ i = j \end{cases} \tag{2.2.10}$$

$$\Longrightarrow \Phi^h(kh) - \sum_{\tau=1}^{k} v(\tau h)\Phi^h\big((k-\tau)h\big) = D^h(kh) \tag{2.2.11}$$

The fact that the matrix $\Phi^h(kh)$is stochastic is already proved in [21,22]. For solving the progression equation proposed an algorithm with suggested matrix form [21]:

$$V^T\Phi^T = D^T \tag{2.2.12}$$

The variables involved are the following:

m= number of states of HSMP, which is 4 in this case.
T = number of periods to be examined for the transient analysis of HSMP.
P = matrix of order m of the embedded Markov chain in HSMP.

G^T = square lower-triangular block matrix order T +1 whose blocks are of order m.
Q^T = kernel of SMP.
Φ^T = block vector of order T + 1 where the blocks are square matrices of order m.
D^T = block vector of order T + 1 where the blocks are the diagonal square matrix of order m.
V^T = square lower-triangular block matrix order T + 1 whose blocks are of order m.
S^T = block vector of order T+1 the block which are the diagonal square matrix of order *m*. The diagonal element of each block *t* are $s_{ii} = \sum_{j=1}^{m} Q_{ij}(t)$.

Given an epoch T is fixed, matrices G and P, the algorithm solves the linear system (2.3.12) for the unknown matrix Φ^T by means of a purely iterative procedure.

3 Results

3.1 Descriptive statistics

This study used data from MoHSS, with 2422 HIV patients on anti-retroviral therapy (ART), who were observed 11028 times (Table 2). Table 1. shows that 1637 (67.6%) were female and 785 (32.41%) were male .657(27.13%) patients started ART in state 1, 683(28.19%) patients started ART in state 2, 677(27.95%) patients started ART in state 3 and 405(16.72%) patients started ART in state 4, at treatment commencement ($t = 0$). Table 2. shows that the highest observation were recorded in the age category of 25-49. Female had the highest observation in all states, except for state 4. Data analysis was done in *msm* (multi-state model) developed by Jackson (2011), the "R package *msm"*, contains numerous functions for fitting continuous-time Markov to longitudinal data. The *msm* package provides several numerical outputs such as sojourn time and transition probabilities.

Table 1. Proportion of male and female patients at the commencement of ART

State	Sex, n (%)		Total
	Male	Female	
1	124 (5.12)	533 (22.01)	657 (27.13)
2	220 (9.08)	463 (19.11)	683 (28.19)
3	246 (10.16)	431 (17.79)	677 (27.95)
4	195 (8.05)	210 (8.67)	405 (16.72)

3.2 Results of semi-Markov for predicting the transitional probabilities

Table 3 shows the estimated transition probability matrix, patient from state 1, 2 and 3 transit to state 4 with probability p<0.001, p<0.001 and 0.018, respectively. Patients show recovery from state 4 to; state 3, state 2 and state 1 with probability of 0.060, 0.002 and p<0.001.

Patients showed recovery from state 3 to 2, from state 3 to 1 and from state 2 to 1 with probability 0.070, 0.003 and 0.071, respectively. The solution of the evolution equation is presented for a specific month in Table 4. It represents the probability that an HIV positive patient being at time 0 in state *i* will be after *t* months, in the state *j*. The conditional probability of a patient starting from state 4 at time zero, and transiting to state 3, 2 and 1 after 2 years is 0.328, 0.227 and 0.162 respectively. A patient being in state 4 at time zero, stayed in same state after 2 years, with 0.288 probability.

The likelihood of direct transition from state 1 to state 2, state 2 to state 3 and state 3 to state 4 after 4 years, are estimated to be 0.284, 0.172 and 0.077 respectively. As *t* increases, the probability of the patient transiting to the next worst state is increasing, while the probability to remain in the same state is decreasing. Table 4 indicates the probability of staying in the same state. The conditional probability that a patient stays in state one, two , three and four for at least 42 months are 0.528, 0.287 , 0.211 and 0.149 respectively. It is decreases with increasing time.

Table 2. Variable description

Variable	State, n = 11028 (%)				Total (n)
	1	2	3	4	
Age*					
<25	135 (50.0)	62 (23.0)	41(15.2)	32(11.9)	270
25-49	3985 (40.3)	2928 (29.6)	2135 (21.6)	837 (8.5)	9885
=>50	253 (29.0)	240 (27.5)	241 (27.6)	139 (15.9)	873
Sex					
Male	887 (25.1)	1105 (31.2)	1037 (29.3)	510 (14.4)	3539
Female	3486 (46.5)	2125 (28.4)	1380 (18.4)	498 (6.6)	7489

Note: n is the number of observations.

Table 3. Estimated transition probability matrix

	State 1	State 2	State 3	State 4
State 1	0.957	0.040	p<0.001	p<0.001
State 2	0.070	0.887	0.041	p<0.001
State 3	0.003	0.071	0.909	0.018
State 4	p<0.001	0.002	0.060	0.937

Table 4. The solution of the evolution equation for month t

Transition	Month=6	Month=12	Month=18	Month=24	Month=30	Month=36	Month=42	Months=48
1→1	0.804	0.698	0.633	0.592	0.563	0.543	0.528	0.518
1→2	0.168	0.233	0.260	0.273	0.278	0.281	0.283	0.284
1→3	0.023	0.059	0.089	0.118	0.128	0.139	0.148	0.154
1→4	0.004	0.010	0.017	0.024	0.030	0.036	0.040	0.044
2→1	0.292	0.405	0.451	0.471	0.479	0.484	0.486	0.486
2→2	0.547	0.387	0.328	0.304	0.293	0.289	0.287	0.286
2→3	0.150	0.183	0.184	0.180	0.177	0.174	0.173	0.172
2→4	0.011	0.026	0.037	0.045	0.049	0.053	0.054	0.056
3→1	0.065	0.168	0.254	0.318	0.363	0.396	0.419	0.436
3→2	0.255	0.309	0.311	0.303	0.295	0.29	0.288	0.286
3→3	0.606	0.425	0.332	0.279	0.247	0.225	0.211	0.201
3→4	0.007	0.098	0.103	0.283	0.094	0.088	0.082	0.077
4→1	0.008	0.045	0.100	0.162	0.220	0.272	0.317	0.353
4→2	0.054	0.130	0.189	0.227	0.251	0.264	0.272	0.277
4→3	0.246	0.327	0.341	0.328	0.306	0.283	0.262	0.243
4→4	0.069	0.497	0.369	0.288	0.222	0.179	0.149	0.127

3.3 Sojourn time

Sojourn time refers to the time a HIV patient spends in each state in a single stay before he/she makes a transition to another state [23]. Table 5 shows estimates of sojourn time, the standard error (SE), the lower bound (L) and the upper bound (U) for each of the transient state *i*. From the results, if an individual is in state 4 he/she spends 15 months in that state, before making transitioning to other states. While a patient spends 22 months in state 1, before transiting to other states.

These two states have the highest sojourn times, mainly because patients in state one have high CD4 count level and the CD4 counts take time to decline. While in state 4, the CD4 count will take time to improve due to time taken by patients, to respond to treatment as state 4 is the worst state, in HIV progression.

Table 5. Sojourn time

	Estimates in months	SE	95%CI
State 1	22.225	0.996	(20.356, 24.266)
State 2	8.133	0.262	(7.636, 8.663)
State 3	10.276	0.415	(9.494, 11.1237)
State 4	15.397	0.958	(13.630, 17.393)

3.4 Prediction of clinical states in individual patient

The probability that a patient starting from state $i \in\{1,2,3,4\}$ at time 0 enters state $j \in\{1,2,3,4\}$ after month t is plotted in Fig. 2. Fig. 2A displays the probability that a patient starting from state 1 at time 0, after month t enters to state $j \in \{1, 2, 3, 4\}$. The probability of remaining in state 1 is higher compared to others, but becomes constant after 75 months.

The probability of a patient starting from state 1 at time zero, enters state $j \in\{1,2,3 ,4\}$ after 108 months, are estimated to be 0.5, 0.4, 0.17 and 0.02 respectively. The conditional probability that a patient starting from state 1 at time zero, enters to state $j \in\{1,2,3,4\}$ after 120 month, are estimated to be 0.5, 0.3, 0.1 and 0.02 respectively.

Fig. 2B shows the probability that a patient starting from state 2 at time 0, after month t enters state $j \in \{1, 2, 3, 4\}$. The probability of remaining in state 2 is high compared to other states for the first 8 months, compared to 9 months. The probability of a patient starting from state 2, at time zero and enters state $j \in \{1,2,3 ,4\}$ after 108 months, are estimated to be 0.5, 0.28, 0.11 and 0.5 respectively. The conditional probability that a patient starting from state 1 at time zero and enters state $j \in \{1,2, 3, 4\}$ after 120 month, is estimated to be 0.5, 0.28, 0.18 and 0.04 respectively.

Fig. 2C shows the probability that a patient starts from state 3 at time 0, after month t and enters stage $j \in \{1, 2, 3, 4\}$. The probability of remaining in state 3 is high compared to other states for the first 20 months, and declined subsequently. The probability of a patient starting from state 3 at time zero and enters state $j \in \{1, 2, 3, 4\}$ after 108 months, is estimated to be 0.5, 0.3, 0.2 and 0.04 respectively. The conditional probability that a patient starting from state 3 at time zero and enters state $j \in\{1,2,3,4\}$ after 120 month, is estimated to be 0.5, 0.3, 0.2 and 0.04 respectively. Fig. 2D shows the probability that a patient starts from state 4 at time 0, after month t and enters stage $j \in \{1, 2, 3, 4\}$. The probability of remaining in state 4 is high, compared to other states for the first 12 months, and declined thereafter until month 85, when it started to increase until month 105, when it started to decrease subsequently. The probability of a patient starting from state 4 at time zero and enters state $j \in \{1, 2, 3, 4\}$ after 108 months, is estimated to be 0.5, 0.29, 0.3 and 0.04 respectively. The conditional probability that a patient starting from state 4, at time zero and enters state $j \in\{1,2,3,4\}$ after 120 months, is estimated to be 0.5, 0.29, 0.3 and 0.04 respectively. The peak influx into State 4 from all other states could be at 100 months, possibly due to adherence to treatment.

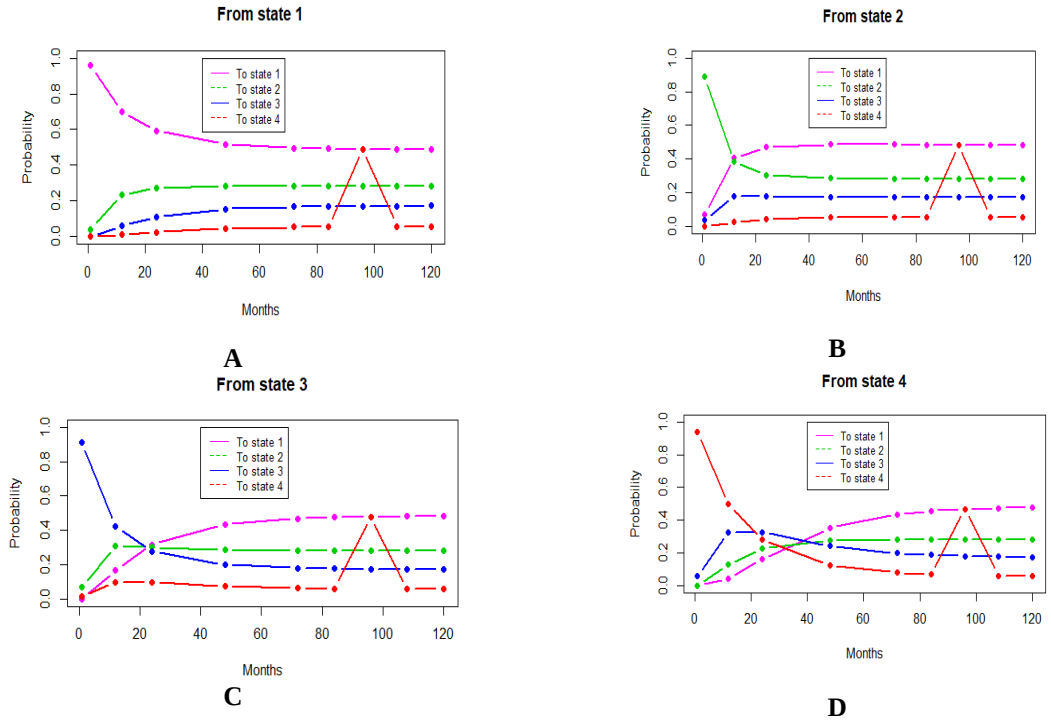

Fig. 2. Conditional probabilities for each state

4 Discussion

It is essential to state that during the follow-up period, no death was recorded and this could be attributed to intensive efforts by the government and various stakeholders. Namibia is the first African country to have reached and exceeded the UNAIDS 2020 goal, to have at least 73% of HIV positive adults viral load suppressed. This slowed down the disease progression and reduced mortality. In terms of 90-90-90 targets, this represents 86% of people with HIV, who knew their status; 96.4% of those on ART and 91.3% of those on treatment, whose viral load was suppressed to <1000 copies/Ml [24].

This paper estimated sojourn times and predicted the transition between clinical states of HIV patients, under ART follow-up in Namibia. Consequently, different plots were produced from the semi-Markov model.

The estimated sojourn times for states 1, 2, 3 and 4 are, 22, 8, 10, and 15 months respectively. If an individual is in state 1, then he/she spends more time in that state, before making a transition to other states. States 1 and 4 have the highest sojourn times, mainly because patients in state one have high CD4 count level and the CD4 counts will take time to decline.

While in state 4 the CD4 count will take time to improve, due to the time it takes for patients to respond to treatment as state 4 is the worst state in HIV progression. Based on these results, policy makers should introduce a policy which compel all patients to be tested for HIV and commence treatment immediately (if they test HIV+), in order to reduce transition probabilities, from state 1, 2 and 3 (good states) to state 4 (worse state). This will also allow patients to spend more time in good states than in worse state.

If an individual is in state 4 then he/she spends more time in that state, before making a transition to other states. This could be due the time taken by an individual, to respond to treatment as state 4 is the worst state, in HIV progression. From a comparable study in South Africa, an author estimated the sojourn time for states one, two, three and four as , 0.88, 0.88, 1.24, 1.20 and 1.57 years respectively [13]. The sojourn time is very important in disease modeling because it state how slowly or fast the disease is progressing.

The conditional probability that a patient starting from state 1 at time zero and enters state $j \in \{1, 2, 3, 4\}$ after 120 months, is estimated to be 0.5, 0.3, 0.1 and 0.02 respectively. The conditional probability that a patient goes from state 1 to 2, from state 2 to 3 and from state 3 to 4 120 months later is 0.3, 0.19 and 0.04 respectively. A similar study in Ethiopia, (Shebeshi), based on data obtained from the antiretroviral therapy unit of Jimma University Specialized Hospital, revealed that the conditional probability that a patient goes from state 1 to 2, from state 2 to 3 and from state 3 to stage 4, occurred 200 months later at 0.27, 0.07 and 0.04 respectively [25].

The probability values are very small; which indicates that as time increases, the conditional probability of transiting to the next worst state is minimal.

The conditional probability that a patient stays in state one, state two, state three and state four after 24 months, are 0.7, 0.39, 0.42 and 0.5 respectively. A similar study by Goshu and Dessie (2013), [6] estimated the probabilities that a patient stays in state 1, 2, 3 and 4 after 24 months are 0.14, 0.0.19, 0.21 and 0.24 respectively. We note that this probability increases with the cumulative and worsening of the illness.

5 Conclusion and Recommendations

Estimating sojourn time and future clinical states is important in understanding HIV progression. The semi-Markov process model is applied, to capture the HIV progression of a patient. The model considers the randomness of the time that a patient spends in a given state of the disease. The sojourn time for state 1, 2, 3 and 4 were estimated. If an HIV patient is in state 1, then he/she spends more time in that state, before making a transition to other states. If an HIV patient is in state 4 then he/she spends more time in that state, before making a transition to other states.

Sojourn time is of interest in HIV modeling, as it gives a signal of how rapidly HIV is progressing. Longer sojourn times, indicates slow HIV progression and shorter sojourn times indicates rapid HIV progression. As time increases, transition probabilities from good states to worse states increases. Without ART, the progression

of HIV will be devastating. It is recommended to stick to ongoing ART treatment, with cautions to patients' recent disease status.

Consent

As per international standard or university standard, patient(s) written consent has been collected and preserved by the author(s).

Disclaimer

This paper is an extended version of a Thesis document of the same author.
The Thesis document is available in this link:
https://repository.unam.edu.na/bitstream/handle/11070/2560/kashihalwa2019.pdf?isAllowed=y&sequence=1

Data Availability Statement

The datasets analyzed during the study are not publicly available, due to patient confidentiality.

Acknowledgements

"This work forms part of a Master study undertaken at the University of Namibia". The authors are thankful to the Ministry of Health and Social Services in Namibia for granting permission and availing data used in this study.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] WHO. Guideline on when to start antiretroviral therapy and on preexposure prophylaxis for HIV. Geneva: WHO Press; 2015.

[2] Laird A. Modeling a progressive disease process under panel observation [PhD thesis]. University of Washington; 2013.

[3] Giuseppe D, DG. A stochastic model for HIV/AIDS dynamic progression. Math Probl Eng. 2007;2007:14.

[4] Viladent C, van Ackere A. HIV/AIDS modeling, a two-angleretrospective. Toward generic deterministic model for pattern II countries? Switzerland: University of Lausanne, Institute of Research in Management. Toward generic deterministic model for pattern II countries?; 2007.

[5] Ross MS. Introduction to probability models. New York: Wiley; 2007.

[6] Goshu AT, Dessie Z. Modelling progression of HIV/AIDS disease stages using semi-markov processes. Data Sci. 2013;11:269-80.

[7] S. Predicting AIDS disease progression using longitudinal CD4 count among adult HIV/AIDS patients in Southwest Ethiopia: application of Semi-Markov processes. Int J Comp Bioinformatics Silico Model. 2016;5(2):808-14.

[8] Masala G, Cannas G, Micocci M. Survival probabilities for HIV infected patients through semi-Markov processes. Biom Lett. 2014;51(1):13-36.
DOI: 10.2478/bile-2014-0002

[9] Zelalem G. Statistical modelling of HIV/ AIDS progression and survival of AIDS patients. A case study of Bahir-Dar [Feleg-Hiwot Referral Hospital MSc thesis]. Hawassa: Hawassa University; 2010.

[10] Petros Kelkile D. Statistical analysis of adult HIV/AIDS patients and modelling of AIDS disease progression. Sci J Appl Math Stat. 2016;4(5):189-201.
DOI: 10.11648/j.sjams.20160405.12

[11] Dessie ZG. Multi-state model of HIV/AIDS by homogenous semi-markov process. Am J Biostat. 2014;4(2):21-8.

[12] Goshu AT, Asena TF. Comparison of sojourn time distributions in modeling HIV/AIDS disease progression. J Biom Biostat. 2017.

[13] Choko C, Chikobua D. Time -homogenous Markov process for HIV/AIDS progression under a combination treatment therapy: cohort study, South Africa. Theor Biol Med Modell. 2018;15(3).

[14] Tamir A. Applications of Markov chains in chemical engineering. Elsevier; 1998.

[15] Barbu VS, Limnios N. Semi-markov chains and hidden semi-markov models toward applications. Springer; 2008.

[16] Mafu TJ. Modelling of multi-state panel data: the importance of the model assumptions; 2014 ([doctoral dissertation]. Stellenbosch: Stellenbosch University).

[17] Moura MDC, Droguett EL. A continuous-time semi-Markov Bayesian belief network model for availability measure estimation of fault tolerant systems. Pesqui Operacional. 2008;28(2):355-75.
DOI: 10.1590/S0101-74382008000200011

[18] Ibe OC. Markov processes for stochastic modelling. Spat Spatio-Temporal Epidemiol; 2009.

[19] Levy P. Processus semi-markoviens. Congress of Mathematicians. Amsterdam. 1954;416-26.

[20] Smith WL. Regenerative stochastic processes. Proc R Soc Lond A. 1955;6-31.

[21] Corradi G, Janssen J, Manca R. Numerical treatment of homogeneous semi-Markov processes in transient case-a straightforward approach. Methodol Comput Appl Probab. 2004;6(2):233-46.
DOI: 10.1023/B:MCAP.0000017715.28371.85

[22] Janssen JN, Manca R. Numerical solution of non-homogenous Semi-Markov processes in transient case. Methodol Comput Appl Probab. 2001;3(3):271-93.
DOI: 10.1023/A:1013719007075

[23] Rubino G, Sericola B. Sojourn times in finite Markov processes. J Appl Probab. 1989;26(4):744-56.
DOI: 10.2307/3214379

[24] Namibia population-based HIV impact assessment NAMPHIA; 2017.

[25] Shebeshi DS. Survival analysis of adult HIV/AIDS patients and stochastic modelling of AIDS disease progression: A case study of [Jimma University MSc thesis]. Addis Ababa: JIMMA University; 2011.

Appendix

Appendix. R codes

```
library(foreign)
library(msm)
library(minqa)
#data=as.data.frame(read.csv("C:\\MSc Biostatistics Thesis\\Katutura HIV data.csv"))
data<-read.table("file:///C:/MSc Biostatistics Thesis/Data/Final HIV data.csv", header=TRUE, sep=",")
P=pmatrix.msm(cav.msm, t = 42, ci = "normal")
PE=round(P$estimates,3)
PL=round(P$L,6)
PU=round(P$U,6)
P
PE
PL
PU
P=pmatrix.msm(cav.msm, t=48, ci = "normal")
S=sojourn.msm(cav.msm)
# predicting future state plot
P=pmatrix.msm(cav.msm, t = 1, ci = "normal")
PE0=P$estimates-P$estimates

for (i in c(1,6,12,18,24,30,36,42,48))
{
P=pmatrix.msm(cav.msm, t = i, ci = "normal")
PE=P$estimates
PE0=cbind(PE0,PE)
}
#PLOTS FROM STATE1 TO STATES
Month=as.data.frame(c(1,6,12,18,24,30,36,42,48))
T1=as.data.frame(PE0[1,c(5,9,13,17,21,25,29,33,37)])
T2=as.data.frame(PE0[1,c(6,10,14,18,22,26,30,34,38)])
T3=as.data.frame(PE0[1,c(7,11,15,19,23,27,31,35,39)])
T4=as.data.frame(PE0[1,c(8,12,16,20,24,28,29,36,40)])

plot(Month$`c(1,6,12,18,24,30,36,42,48)`,
    T1$`PE0[1, c(5, 9, 13, 17, 21, 25, 29, 33, 37)]`,
    type = "b",col=6,lwd=2,ylim=c(0,1),pch=16,xlab="Months",ylab="Probability",
    main = "From state 1")

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
    T2$`PE0[1, c(6, 10, 14, 18, 22, 26, 30, 34, 38)]`,
    type = "b",col=3,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T3$`PE0[1, c(7, 11, 15, 19, 23, 27, 31, 35, 39)]`,
     type = "b",col=4,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T4$`PE0[1, c(8, 12, 16, 20, 24, 28, 29, 36, 40)]`,
     type = "b",col=2,lwd=2,pch=16)

legend(40, 1, legend=c("To state 1", "To state 2","To state 3","To state 4"),
      col=c(6,3,4,2), lty=1:2, cex=0.8)
#PLOTS FROM STATE2 TO STATES
Month=as.data.frame(c(1,6,12,18,24,30,36,42,48))
```

```
T1=as.data.frame(PE0[2,c(5,9,13,17,21,25,29,33,37)])
T2=as.data.frame(PE0[2,c(6,10,14,18,22,26,30,34,38)])
T3=as.data.frame(PE0[2,c(7,11,15,19,23,27,31,35,39)])
T4=as.data.frame(PE0[2,c(8,12,16,20,24,28,29,36,40)])

plot(Month$`c(1,6,12,18,24,30,36,42,48)`,
    T1$`PE0[2, c(5, 9, 13, 17, 21, 25, 29, 33, 37)]`,
    type = "b",col=6,lwd=2,ylim=c(0,1),pch=16,xlab="Months",ylab="Probability",
    main = "From state 2")
lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T2$`PE0[2, c(6, 10, 14, 18, 22, 26, 30, 34, 38)]`,
     type = "b",col=3,lwd=2,pch=16)
lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T3$`PE0[2, c(7, 11, 15, 19, 23, 27, 31, 35, 39)]`,
     type = "b",col=4,lwd=2,pch=16)
lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T4$`PE0[2, c(8, 12, 16, 20, 24, 28, 29, 36, 40)]`,
     type = "b",col=2,lwd=2,pch=16)
legend(40, 1, legend=c("To state 1", "To state 2","To state 3","To state 4"),
      col=c(6,3,4,2), lty=1:2, cex=0.8)
#PLOTS FROM STATE3 TO STATES
Month=as.data.frame(c(1,6,12,18,24,30,36,42,48))
T1=as.data.frame(PE0[3,c(5,9,13,17,21,25,29,33,37)])
T2=as.data.frame(PE0[3,c(6,10,14,18,22,26,30,34,38)])
T3=as.data.frame(PE0[3,c(7,11,15,19,23,27,31,35,39)])
T4=as.data.frame(PE0[3,c(8,12,16,20,24,28,29,36,40)])

plot(Month$`c(1,6,12,18,24,30,36,42,48)`,
    T1$`PE0[3, c(5, 9, 13, 17, 21, 25, 29, 33, 37)]`,
    type = "b",col=6,lwd=2,ylim=c(0,1),pch=16,xlab="Months",ylab="Probability",
    main = "From state 3")

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T2$`PE0[3, c(6, 10, 14, 18, 22, 26, 30, 34, 38)]`,
     type = "b",col=3,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T3$`PE0[3, c(7, 11, 15, 19, 23, 27, 31, 35, 39)]`,
     type = "b",col=4,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T4$`PE0[3, c(8, 12, 16, 20, 24, 28, 29, 36, 40)]`,
     type = "b",col=2,lwd=2,pch=16)

legend(40, 1, legend=c("To state 1", "To state 2","To state 3","To state 4"),
      col=c(6,3,4,2), lty=1:2, cex=0.8)


#PLOTS FROM STATE4 TO STATES
Month=as.data.frame(c(1,6,12,18,24,30,36,42,48))

T1=as.data.frame(PE0[4,c(5,9,13,17,21,25,29,33,37)])
T2=as.data.frame(PE0[4,c(6,10,14,18,22,26,30,34,38)])
T3=as.data.frame(PE0[4,c(7,11,15,19,23,27,31,35,39)])
T4=as.data.frame(PE0[4,c(8,12,16,20,24,28,29,36,40)])
```

```
plot(Month$`c(1,6,12,18,24,30,36,42,48)`,

    T1$`PE0[4, c(5, 9, 13, 17, 21, 25, 29, 33, 37)]`,
    type = "b",col=6,lwd=2,ylim=c(0,1),pch=16,xlab="Months",ylab="Probability",
    main = "From state 4")

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T2$`PE0[4, c(6, 10, 14, 18, 22, 26, 30, 34, 38)]`,
     type = "b",col=3,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T3$`PE0[4, c(7, 11, 15, 19, 23, 27, 31, 35, 39)]`,
     type = "b",col=4,lwd=2,pch=16)

lines(Month$`c(1,6,12,18,24,30,36,42,48)`,
     T4$`PE0[4, c(8, 12, 16, 20, 24, 28, 29, 36, 40)]`,
     type = "b",col=2,lwd=2,pch=16)

legend(40, 1, legend=c("To state 1", "To state 2","To state 3","To state 4"),
      col=c(6,3,4,2), lty=1:2, cex=0.8)
```


Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/97022