Estimating sojourn time and transition between clinical states of HIV patients under ART follow up in Namibia

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 semi-Markov model was employed to estimate sojourn times and transition between clinical states.
Results: As expected the probabilities of transiting from good states 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 increase the probabilities of remaining 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: HIV, sojourn time, semi-Markov processes, multi-state model, ART

1. INTRODUCTION

HIV disease is one of the leading causes of death in Namibia and worldwide [1]. HIV disease does not only have massive economic impact through lost productivity and medical care expenditure, but it is also one of the main reason of disability and human su□ering [2]. 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 that can be in any states and that each time it makes a transition to any state; it remains there for a random amount of time and then makes a forward or backward transition into another state with some probability [5].

Recent studies on HIV have estimated sojourn times and transition between clinical states using homogenous semi-Markov processes. Dynamical models to estimate the proportion of individuals changing their status at each time step were used in many studies [6-11]. Semi-Markov models were applied to HIV disease evolution and compared sojourn time distributions, exponential and Weibull probability distributions [12]. More over other authors applied homogenous Markov process to HIV disease under a combination treatment therapy [13].

2. MATERIAL 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 are HIV infected and whose CD4 counts were measured at least once were included in this study. A total number of 2422 patients who were observed 11028 times constituted the study. 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.

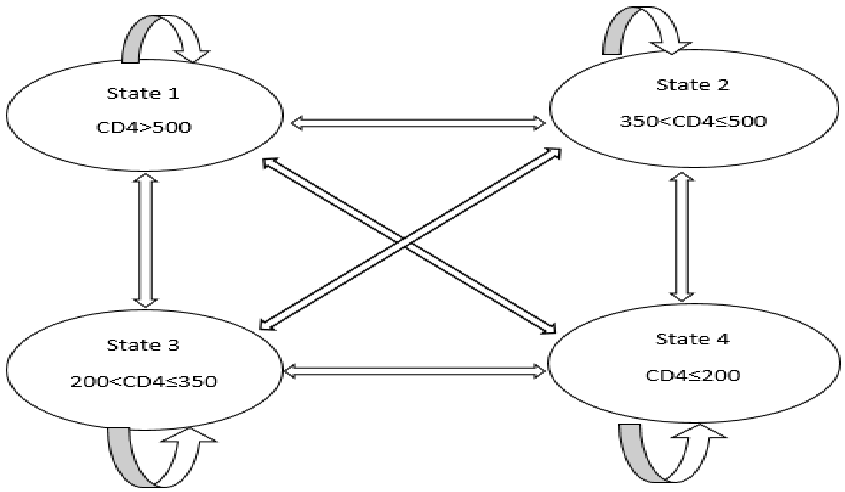


Fig. 1. Transition diagram

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 assumption is that the 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 a process 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 $S = \{S_1, S_2, \ldots, S_m\}$ 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. Here 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:

$$Q_{ij}(t) = P[T_{n+1} = j, T_{n+1} - T_n \leq t | X_n = i] \quad (2.3.1)$$

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

$$P_{ij} = \lim_{t \to \infty} Q_{ij}(t) \quad (2.3.2)$$

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

$$H_i(t) = P[T_{n+1} - T_n \le t | X_n = i] = \sum_{j=1}^{m} Q_{ij}(t) \tag{2.3.3}$$

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

$$G_{ij}(t) = P[T_{n+1} - T_n \le t | X_n = i, X_{n+1} = j], \tag{2.3.4}$$

which can be computed as:

$$G_{ij}(t) = \begin{cases} \frac{Q_{ij(t)}}{P_{ij}}, & if P_{ij} \neq 0 \\ 1, & if P_{ij} = 0 \end{cases} \tag{2.3.5}$$

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).

$$\emptyset_{(ij)}(t) = P[X(t) = j | X(0) = i], \tag{2.3.6}$$

$$\emptyset_{ij}(t) = (1 - H_i(t))\delta_{ij} + \sum_{l=1}^{m} \int_0^t Q_{il}(\tau)\phi_{lj}(t-\tau)\, d\tau \tag{2.3.7}$$

Here 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:

$$\phi_{ij}^{h}(kh) = d_{ij}^{h}(kh) + \sum_{i=1}^{m} \sum_{\tau=1}^{k} v_{ij}^{h}(\tau h)\phi_{ij}^{h}((k-\tau)h) \tag{2.3.8}$$

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

$$d_{ij}^{h}(kh) = \begin{cases} 0 & if\ i \neq j, \\ 1 - H_i^h(kh), & if\ i = j, \end{cases} \tag{2.3.9}$$

$$v_{ij}^{h}(kh) = \begin{cases} 0, & if\ i \neq j \\ \varrho_{ij}^{h}(kh) - \varrho_{ij}^{h}((k-1)h), & if\ i = j \end{cases} \tag{2.3.10}$$

$$\Longrightarrow \Phi^h(kh) - \sum_{\tau=1}^{k} v(\tau h)\Phi^h((k-\tau)h) = D^h(kh) \tag{2.3.11}$$

The fact that the matrix $\Phi^h(kh)$is stochastic is already proved in [21, 22]. For solving the progression equation, Corradi *et al*. (2004) [21] proposed an algorithm with suggested matrix form:

$$V^T\Phi^T = D^T \tag{2.3.12}$$

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 AND DISCUSSION

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 females and 785 (32.41%) were males , 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 have 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

	Sex, n (%)		
State	**Male**	**Female**	**Total**
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 show 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 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, stay in same state after 2 years with probability 0.288.

The probabilities 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 a next worse state is increasing while the probability to remain in the same state is decreasing. **Table 4** also 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 decreasing with increasing time.

Table 2. Variable description

	State, n = 11028 (%)				
Variable	**1**	**2**	**3**	**4**	**Total (n)**
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 a transition 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 will 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 since 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)

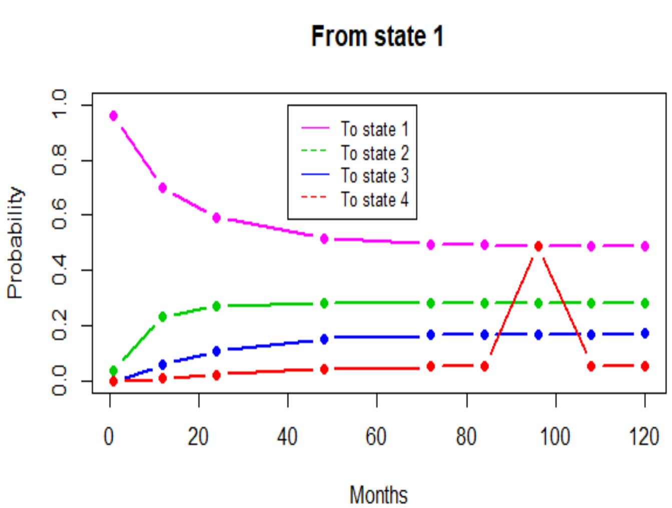
3.4 Prediction of clinical states in individual patient

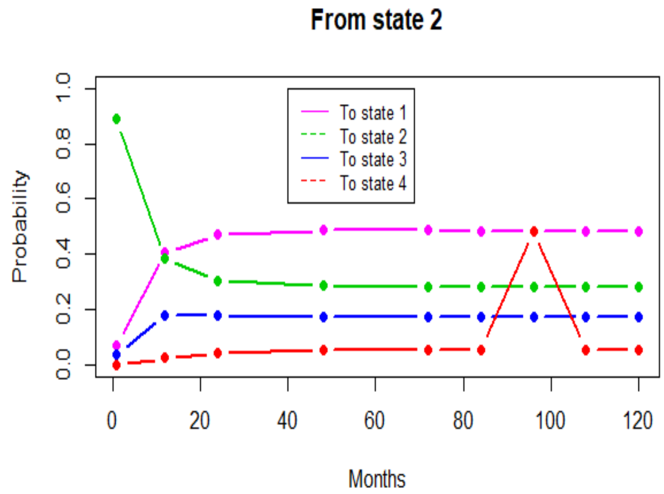
The probability that a patient starting from state i $\in\{1,2,3,4\}$ at time 0 enters state $j \in\{1,2,3,4\}$ after month t is plotted in **Fig. 2**. **Fig. 2A** shows the probability that a patient starting from state 1 at time 0, after month t enters to state j $\in$ {1, 2, 3, 4}. The probability of remaining in state 1 is high as 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 as compared to others for the first 8 month, then after 9 month. The probability of a patient starting from state 2 at time zero 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, enters state $j \in \{1,2, 3, 4\}$ after 120 month, are estimated to be 0.5, 0.28, 0.18 and 0.04 respectively.

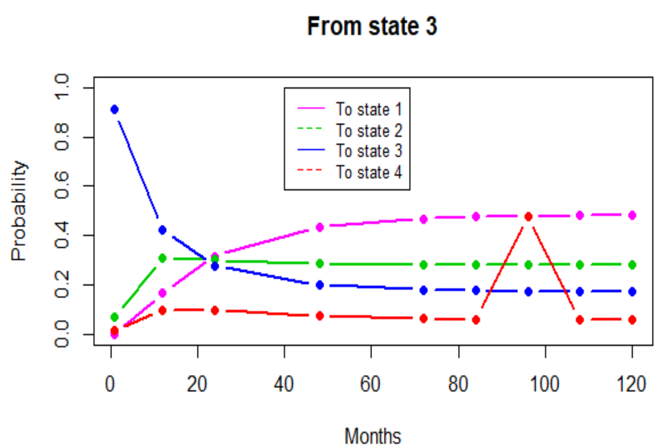
Fig. 2C shows the probability that a patient starting from state 3 at time 0, after month t enters to stage $j \in \{1, 2, 3, 4\}$. The probability of remaining in state 3 is high as compared to others for the first 20 month, and declined subsequently. The probability of a patient starting from state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

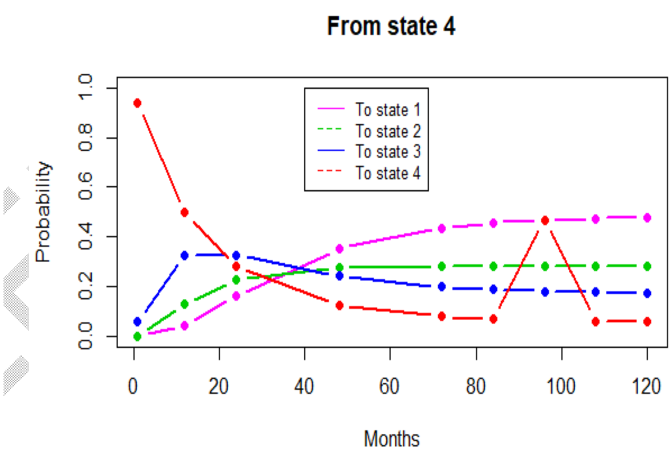




B

A





C **D**

Fig. 2. Conditional probabilities for each state

4. DISCUSSION

It is essential to state that during the follow-up period no death was recorded, this could be attributed to concerted efforts by 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 virally suppressed, this slow down the disease progression and reduce mortality. In terms of 90-90-90 targets, this represents 86% of people with HIV who reported knowing their status; 96.4% of those on ART; and 91.3% of those treated virally suppressed to <1000 copies/Ml [24].

This paper estimated sojourn times and predicted the transition between clinical states of HIV patient 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 month 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 time taken by patients to respond to treatment since state 4 is the worst state in HIV progression. Based on these results policy makers should pass a policy which compel all sick individual to be tested for HIV and commence treatment immediately(if they test HIV+) so as to reduce transition probabilities from good states(state 1, 2 and 3) to worse state (state 4). 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 since 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 as it tell how slowly or fast the disease is progressing.

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. The conditional probability that a patient goes from state 1 to 2, from state 2 to 3 and from state 3 to 4 120 month 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 4 200 months later is 0.27, 0.07 and 0.04 respectively [25].

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

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, Goshu and Dessie. (2013) [6] estimated the probabilities that a patient stays in state 1, 2, 3 and 4 after 24 months , 0.14, 0.0.19, 0.21 and 0.24 respectively. We note that this probability is increasing with the increasing seriousness of the illness.

4.1 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 a HIV patient is in state 1 then he/she spends more time in that state before making a transition to other states. If a 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 quickly 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.

DATA AVAILABILITY STATEMENT

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

ETHICAL APPROVAL (WHERE EVER APPLICABLE)

REFERENCES

[1] WHO. (2015). *Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV.* Geneva: WHO Press.
[2] Laird A. (2013). Modeling a Progressive Disease Process Under Panel Observation. PhD Thesis, University of Washington.
[3] Giuseppe D, D. G. (2007). A stochastic model for HIV/AIDS dynamic progression. *Mathematical Problems in Engineering, 2007*, 14.

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

[5] Ross M S. (2007). *Introduction to Probability Models.* New York: Wiley.

[6] Goshu A T & Dessie Z. (2013). Modelling Progression of HIV/AIDS Disease Stages Using Semi-Markov Processes. *Data Science, 11(2013)*, 269-280.

[7] S. (2016). Predicting AIDS disease progression using longitudinal CD4 count among adult HIV/AIDS patients in Southwest Ethiopia: Application of Semi-Markov processes. *International Journal of Computational Bioinformatics and In Silico Modeling, 5 No 2(2016)*, 808-814.

[8] Masala G, Cannas G and Micocci M. (2014). Survival probabilities for HIV infected patients through semi-Markov processes. *Biometrical Letters, 51(2014)*.

[9] Zelalem G. (2010). *Statistical Modelling of HIV/ AIDS Progression and Survival of AIDS Patients. A Case Study of Bahir-Dar, Feleg-Hiwot Referral Hospital.* M.Sc. Thesis, Hawassa University, Hawassa.

[10] Kelkile P D. (2016). Statistical Analysis of Adult HIV/AIDS Patients and Modelling of AIDS Disease Progression. *Science Journal of Applied Mathematics and Statistics, 4(5)*, 189-201.

[11] Dessie ZG. (2014). Multi-state model of HIV/AIDS by homogenous semi-markov process. American Journal of Biostatistics, 4(2), 21-28.

[12] Goshu A T and Asena T F. (2017). Comparison of Sojourn Time Distributions in Modeling HIV/AIDS Disease Progression. *Journal of Biometrics & Biostatistics*.

[13] Choko C and Chikobua D. (2018). Time -homogenous Markov process for HIV/AIDS progression under a combination treatment therapy: Cohort study, South Africa. *Theoretical Biology and Medical modelling, 15*(3).

[14] Tamir, A. (1998). *Applications of Markov chains in chemical engineering*. Elsevier.

[15] Barbu V S & Limnios N. (2008). Semi-Markov Chains and Hidden Semi-Markov Models toward Applications. Springer.

[16] Mafu, T. J. (2014). *Modelling of multi-state panel data: the importance of the model assumptions* (Doctoral dissertation, Stellenbosch: Stellenbosch University).

[17] Moura, M. D. C., & Droguett, E. L. (2008). A continuous-time semi-Markov Bayesian belief network model for availability measure estimation of fault tolerant systems. *Pesquisa Operacional*, *28*(2), 355-375.

[18] Ibe, O.C. (2009). Markov processes for stochastic modelling. *Spatial and Spatio-temporal Epidemiology*.

[19] Levy P. (1954). Processus semi-Markoviens. *Congress of Mathematicians*, (pp. 416-426). Amsterdam.

[20] Smith W.L. (1955). Regenerative stochastic processes. *Proceedings of the Royal Society of London Series A*, (pp. 6-31). London.

[21] Corradi, G., Janssen, R. and Monica, R. (2004). Numerical treatment of homogeneous semi-Markov processes in transient case-a straightforward approach. *Methodology and Computing in Applied Probability* 6, 233-246.

[22] Jansen JN & Monica. (2001). Numerical solution of non-homogenous Semi-Markov processes in transient case. Methodology and Computing in Applied Probability, 3, 271-279.

[23] Rubino, G., and Sericola, B. (1989). Sojourn times in finite Markov processes. *Journal of Applied Probability*, 26(4), 744-756.

[24] NAMIBIA POPULATION-BASED HIV IMPACT ASSESSMENT NAMPHIA 2017

[25] Shebeshi DS. (2011). *SURVIVAL ANALYSIS OF ADULT HIV/AIDS PATIENTS AND STOCHASTIC MODELLING OF AIDS DISEASE PROGRESSION: A CASE STUDY OF JIMMA UNIVERSITY.* MSc Thesis, JIMMA University, Addis Ababa.

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)
```