Semi-Markov model for evaluating the HIV patient treatment cost

Waema R. Mbogo[1], Livingstone S. Luboobi[2] and John W. Odhiambo[3]

[1] [3]Center for Applied Research in Mathematical Sciences
Strathmore University , Box 59857 00200, Nairobi - KENYA
rmbogo@strathmore.edu

[2]Department of Mathematics
Makerere University , Box 7062, Kampala - UGANDA
luboobi@math.mak.ac.ug

Abstract

The aim of this study is to model the progression of HIV/AIDS disease and evaluate the cost of the anti-retroviral therapy for an HIV infected patient under ART follow-up using Non homogeneous semi-Markov processes. States of the Markov process are defined by the seriousness of the sickness based on the clinical scores. The five states considered are: Asymptomatic (CD_4^+ count > 500 cells/microliter); Symptomatic 1 ($350 < CD_4^+$ count ≤ 500 cells/microliter); Symptomatic 2 ($200 < CD_4^+$ count ≤ 350 cells/microliter); AIDS (CD_4^+ count ≤ 200 cells/microliter) and Death (Absorbing state). The first four states are named as good or alive states.

The models formulated can be used to gain insights on the transition dynamics of the HIV patient given the follow-up time. The transition probability Model, when fitted with data will give insights on the conditional probability of a patient moving from one disease state to another, given the current state and the follow-up time. This model will also give the probability of survival for the HIV patient under treatment given the current state and follow-up time.

---

[1]corresponding author's email: rmbogo@strathmore.edu (Rachel W. Mbogo)

The total Lifetime Treatment Cost model obtained, when applied to real data will give the cost of managing an HIV patient given the starting state, the treatment combination which incurs minimum cost and which treatment combination is most effective at each state. The treatment reward model also when applied to real data will give the state, which a patient should be maintained so that they remain healthy, non-infectious and productive to the society. Also the model will show the optimal/effective time to initiate treatment, which can be used to give advice on how to handle the HIV infecteds given their states.

Keywords: HIV evolution, Therapeutic intervention, Non homogeneity, Treatment Cost, Semi-Markov model, transition probabilities, reward model.

1   Introduction

As extensions of Markov processes and renewal processes, semi-Markov processes are widely applied and hence, an important methodology for modeling. Semi-Markov models have extensively been studied and applied in finance, insurance, business administration as well as manpower models. In biology and medicine, Semi-Markov modeling has also been used in continuous time to study prognosis and the evolution of diseases, see [22]; [18]; [8]; [21]; [19]; [15] [1]; [9]; [16] ;[5]; [10]; [20]; [17]. Typically these methods assume the sample paths are continuously observed. However, it is often the case where study individuals' states are observed only at discrete time points with no information about the types and times of events between observation times.

Recently, [17] developed methods for fitting continuous-time semi- Markov multi-state models to panel data. Their methods are illustrated with a model of the natural history of oncogenic genital HPV infection in women using data from the placebo arm of an HPV vaccine trial. Discrete-time semi-Markov models have not received as much attention in the literature as continuous-time semi- Markov models. In finance, for example, credit rating and reliability models are based upon discrete time Semi-Markov theory like [12]; [24]; [25]; [23]; [2] ;[3]; [7]. For studies with fixed scheduled visits, such as clinical trials, it is natural to model time as discrete.

Discrete-time models can have advantages over continuous-time models, such as not requiring the specification of guarantee times. [12];[23] and [25] studied nonparametric estimators for discrete-time semi-Markov unidirectional models with varying initial states in HIV data. They considered only unidirectional models, which may not be applied to complex disease processes such as HPV

and HIV where prior states may be revisited and states may not be visited sequentially. Their methods extended the Markov models developed by [6]; [11] and [13] to a more generalizable discrete-time semi-Markov framework by allowing the probability of transitioning from the HIV positive state to AIDS to depend on the duration of HIV infection.

[2] and [3] studied discrete-time multi-state bidirectional semi-Markov models in finance. Their methods require parametric assumptions, only allow for incident infections, and do not address the possibility of missing data.

The evolution of HIV is Multi-directional, that is recovery from one state to previous state is reasonable. We propose to develop discrete-time non-homogeneous Semi- Markov models that can be used to study HIV evolution.

2 HIV progression Dynamics

Disease progression models are useful tools for gaining a systems' understanding of the transitions to disease states, and characterizing the relationship between disease progress and factors affecting it such as patients' profile, treatment and the HIV diagnosis stage. The natural history of HIV infection can be considered as a series of stages through which a patient progresses. Based both on current information and physicians opinion, we have taken 5 state classifications as categorized by clinical signs. In order to predict the HIV evolution, we assume the disease progresses following four transient states related to clinical scores plus an absorbing state (the death of the patient). Transitions are allowed in all the transient states. The absorbing state is categorized as "bad" (once a patient is in the death state, she/he will never be in the other states and rather stays there forever) and the transient states as "good" states (the patient is alive).

1. State I : Asymptomatic (CD_4^+ $count > 500$ cells/microliter)

2. State II : Symptomatic 1 ($350 < CD_4^+$ $count \leq 500$ cells/microliter).

3. State III : Symptomatic 2 ($200 < CD_4^+$ $count \leq 350$ cells/microliter).

4. State IV : AIDS (CD_4^+ $count \leq 200$ cells/microliter).

5. D : Death (Absorbing state).

With the above states, we have the HIV transition diagram as shown in Figure 1.
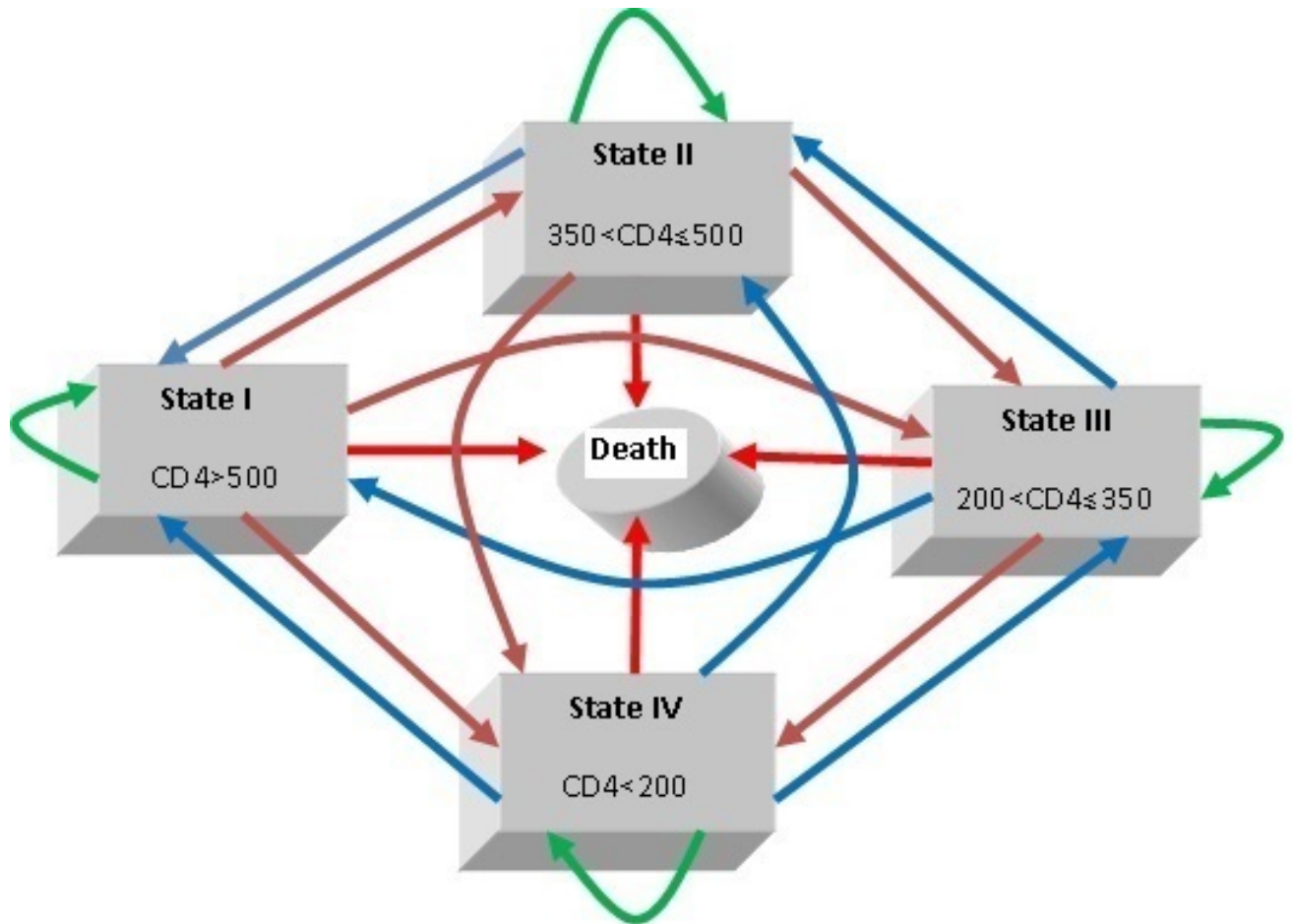
Figure 1: HIV Multi-state Model with 5 immunological states and 20 transitions

The red arrows show the progression of a patient to a worse state, blue arrows show recovery of a patient to a better state and green arrows show a patient remaining in the same state.

From the transition diagram in Figure 1, we have HIV transition matrix as shown in Table 1.

Table 1: Classification of HIV disease State as defined by CDC/WHO

Classification of HIV Disease States					
Clinical scores	1: Asymptomatic	2: symptomatic 1	3: Symptomatic 2	4: AIDS	D:Death
1: Asymptomatic	n_{11}	n_{12}	n_{13}	n_{14}	n_{1D}
2: symptomatic 1	n_{21}	n_{22}	n_{23}	n_{24}	n_{2D}
3: Symptomatic 2	n_{31}	n_{32}	n_{33}	n_{34}	n_{3D}
4: AIDS	n_{41}	n_{42}	n_{43}	n_{44}	n_{4D}
D:Death	n_{D1}	n_{D2}	n_{D3}	n_{D4}	n_{DD}

The entries n_{ij} in the table represents the number of HIV patients who have moved from state i to state j, where i and j are the disease states as defined by CDC [4]. The transition $n_{Dj} \quad j = 1, 2, 3, 4$ is not possible because once a patient is in the death state, she/he will never be in the other states and rather stays there forever.

For the clinical classification: 1 represents - Asymptomatic (No HIV related symptoms seen), 2 - represents Symptomatic-Moderate unexplained weight loss ($< 10\%$ of the body weight), 3 - represents Symptomatic-Unexplained severe weight loss ($> 10\%$ of the body weight) and 4 - represents AIDS-the patient has other co-infections.

Note: Patients in Category 4 are considered to have AIDS.

We model the progression of an HIV patient in different disease states as a Discrete-time non-homogeneous semi-Markov stochastic process $\{X(t), t \geq 0\}$, having five possible states constituting the state space $E = \{I, II, III, IV, D\}$, defined on the probability space (Ω, Γ, P). All the states apart from D are inter-related, which means improvements are also considered. Patients move through these five states according to twenty transitions as illustrated in the state transition diagram in Figure 1.

Now we define the following random variables as defined by [14]

$X_n : \Omega \to E, T_n : \Omega \to N; n \geq 0$

Where X_n represents the state of the system at the nth transition. T_n represents the chronological time of the nth transition. A sequence $\{X_n\}_{n \geq 0}$ of

random variables with values in the set E, represents the state of E at time n. If $X_n = i$, the process is said to be in state i at time n, or to visit state i at time n.

Let us define the *duration process* $(S_n)_{n \in N}$ by

$$
\begin{aligned}
S_0 &= 0 \\
S_{n+1} &= T_{n+1} - T_n
\end{aligned}
$$

where S_{n+1} represents the duration time spent in state X_n. The $(X_n, T_n)_{n \in N}$ process is called the non-homogeneous Markov renewal process if:

$$
\begin{aligned}
Q_{ij}(t) &= P[X_{n+1} = j, S_{n+1} \leq t | X_n = i, S_n = s, X_{n-1}, S_{n-1}, ...X_1, S_1, X_0, S_0] \\
&= P[X_{n+1} = j, S_{n+1} \leq t | X_n = i]
\end{aligned}
$$

and for $j \neq i$

The term

$$
Q_{ij}(t) = P[X_{n+1} = j, S_{n+1} \leq t | X_n = i]
$$

is the associated non-homogeneous semi-markov kernel Q, which is the conditional probability that the process is at state j at time t, given that the process was in state i at $t = 0$ and remained in state i for a period of $S_{n+1} \leq t$ time units.

3   HIV Non-Homogeneous Semi-Markov Model

In order to consider the effects due to medical scientific progress of the HIV patient, we consider the following sequences of random variables:

$X_n : \Omega \to E, T_n : \Omega \to N; n \in N$

where X_n represents the state at the nth transition, that is the possible state in which the infection may show its level of seriousness, with the set of disease state $E = \{I, II, III, IV, D\}$. T_n represents the chronological time in which the nth transition occurred. A sample path of the disease evolution is shown in Figure 2.
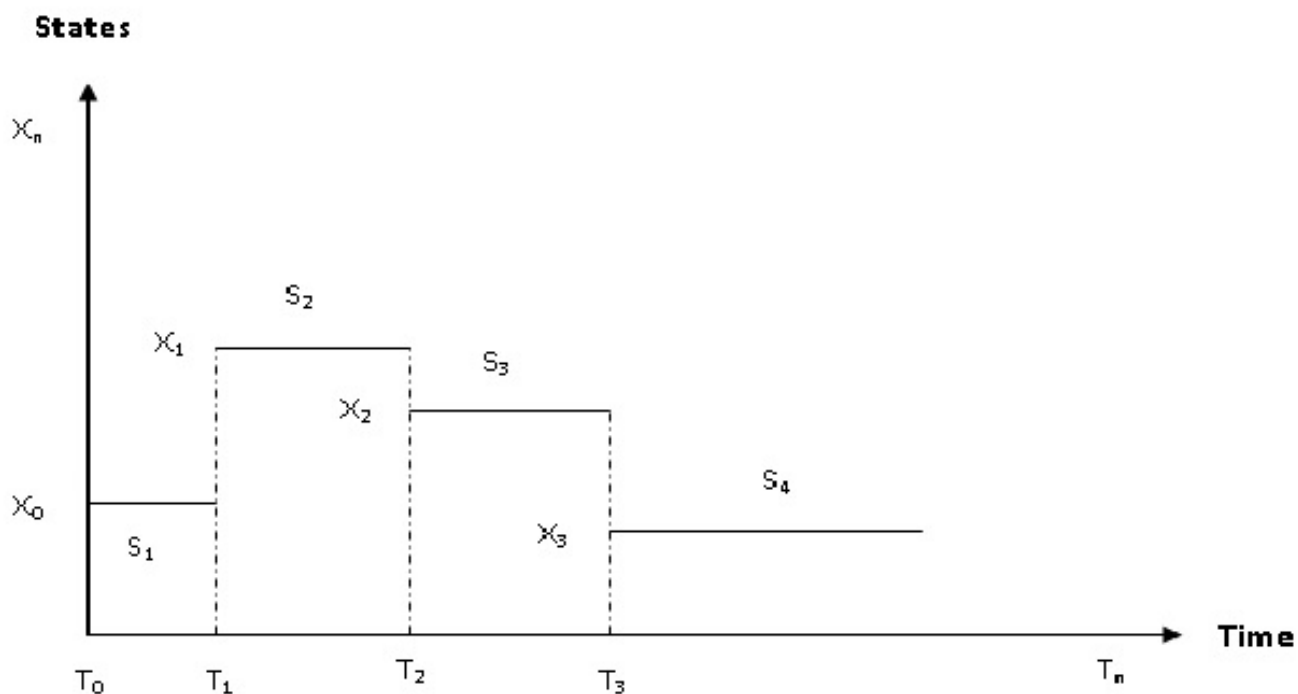
Figure 2: Trajectory of a semi-Markov process

X_n represents the disease state, T_n represents the disease observation time, and S_n is the sojourn time (holding time) with S_{n+1} representing the duration time spent in state X_n;

where

$$S_0 = 0$$
$$S_{n+1} = T_{n+1} - T_n$$

It is supposed that the process (X_n, T_n) is a non homogeneous Markovian renewal process with kernel $Q = [Q_{ij}(x, t)]$ defined in the following way:

$$Q_{ij}(x, t) = P[J_{n+1} = j, S_{n+1} \leq t | J_n = i, S_n = x]$$

which represents the conditional probability that a patient is in state j of the disease at the $(n + 1)^{th}$ transition within the chronological time t, given that she/he entered state i of the disease at time x with the nth transition.

Let $p_{ij}(x, t)$ represent the probability that a patient makes his/her next transition to state j at time t, given that he entered state i at time x. Therefore $\mathbf{P}(t) = [p_{ij}(t)]_{i,j}$ is the $(5 * 5)$ transition probability matrix of the embedded non-homogeneous Markov chain $(X_n)_{n \in N}$ and must satisfy $\int_0^\infty \sum_j p_{ij}(t)dt = 1$. The summation equation expresses the requirement that there is unit probability that an HIV patient will be in one of the five states of the disease progression at some point in the future, given it started in state i. Hence

$$p_{ij}(x) = P[X_{n+1} = j | X_n = i] = \lim_{t \to \infty} Q_{ij}(x, t) \quad 1 \le i, j \le 5$$

From the transition matrix in Table 1, we have the probability transition matrix

$$\lim_{t \to \infty} P(t) = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{1D} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{2D} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{3D} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{4D} \\ p_{D1} & p_{D2} & p_{D3} & p_{D4} & p_{DD} \end{pmatrix}$$

When a patient makes a transition from state i to state j, the transition is said to be real if $i \ne j$, otherwise it is virtual, that is when $i = j$. Virtual transitions are represented in transition matrix by non-zero diagonal elements. However, before the entrance into j, the patients holds for a time t in state i. The governing equation for these probabilities are derived as follows:-

A system starting in state i can be in state j at time t in the following ways

1. $i = j$ and the system never leaves state i or the system leaves state i but returns to state i by time t.

2. $i \ne j$ and the system leaves state i and manages to reach state j by time t.

Let w_i be the mean sojourn time in state i, and let $p_{ij}(x)$ be the probability that a patient makes his/her next transition to state j, given that he entered state i at time x. With the transition probability matrix P and the sojourn time vector w, we can visualize the evolution of the semi-Markov process as follows:

A patient starts in state i, stays there on average for w_i amount of time and then progresses to state j with probability $p_{ij}(x)$, stays in state j for w_j amount of time on average and then moves to another state and so on.

Therefore the conditional probability of a patient being at state j at time t, given that the patient was in state i at x and remained in state i for a period of w_i time units, can be represented as

$$
\begin{aligned}
Q_{ij}(x,t) &= P[X_{n+1} = j, S_{n+1} \leq t | X_n = i, S_n = x] \\
&= w_i(x,t) p_{ij}(x,t)
\end{aligned}
\tag{1}
$$

The probability that a transition takes place from any state in time t is given by summing up all the leaving probabilities from state i to each possible state j. That is, the probability that a patient moves from state i to any other state within time t, given some waiting time in state i is given by:

$$
\begin{aligned}
Q_i(x,t) &= P[S_{n+1} \leq t | X_n = i, X_{n+1} = j, S_n = x] \\
&= \sum_{j \neq i}^{4} Q_{ij}(x,t) \\
&= \sum_{j \neq i}^{4} w_i(x,t) p_{ij}(x,t)
\end{aligned}
\tag{2}
$$

Therefore, the probability that the patient does not leave state i in time t is given by

$$
H_i(x,t) = P\{S_{n+1} = t | X_n = i, S_n = x\} = 1 - Q_i(x,t)
\tag{3}
$$

Equation (3) is the conditional probability that a patient is still in state i at time t, given that the patient entered state i at time x.

Combining equations 1, 2 and 3 we get the equation that describes the transition probabilities which include sojourn times in the states, that is:

$$
\begin{aligned}
p_{ij}(x,t) &= \delta_{ij}[H_i(x,t)] + \sum_{k \neq i}^{4} \int_x^t Q_{ik}(\tau,t) p_{kj}(x,\tau) d\tau \\
&= \delta_{ij}[H_i(x,t)] + \sum_{k \neq i}^{4} \int_x^t Q_{ik}(x,\tau) p_{kj}(\tau,t) d\tau \\
&= \delta_{ij}[1 - Q_i(x,t)] + \sum_{k \neq i}^{4} \int_x^t w_i(x,\tau) p_{ik}(x,\tau) p_{kj}(\tau,t) d\tau
\end{aligned}
\tag{4}
$$

where δ_{ij} represents the Kronecker delta function defined by:-

$$\delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

When i is an absorbing state, then

$$Q_{ij}(x,t) = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Therefore

$$p_{ij}(x,t) = \delta_{ij} \tag{5}$$

When i is transient and j is an absorbing state, then

$$p_{ij}(t) = Q_{ij}(x,t) + \sum_{k \neq i}^{4} \int_{x}^{t} w_i(x,\tau)p_{ik}(x,\tau)p_{kj}(\tau,t)d\tau \tag{6}$$

The time in equation (4) is continuous, which is applicable in laboratory situation. For HIV disease, we assume that the observational time is discrete , therefore we model the evolution as Discrete time non-homogeneous Semi-Markov process. The discrete time semi-Markov model is given by

$$\begin{aligned}
p_{ij}(x,t) &= \delta_{ij}[H_i(x,t)] + \sum_{k \neq i}^{4} \sum_{\tau=0}^{t} Q_{ik}(x+\tau, t-\tau)p_{kj}(x,\tau) \\
&= \delta_{ij}[H_i(x,t)] + \sum_{k \neq i}^{4} \sum_{\tau=0}^{t} Q_{ik}(x,\tau)p_{kj}(x+\tau, t-\tau) \tag{7} \\
&= \delta_{ij}[1 - Q_i(x,t)] + \sum_{k \neq i}^{4} \sum_{\tau=0}^{t} w_i(x,\tau)p_{ik}(x,\tau)p_{kj}(x+\tau, t-\tau)
\end{aligned}$$

The first term on the right hand side of equation (7) is the probability that an HIV patient being in state i never leaves state i until the end of the period t. In this case $i = j$ and $\delta_{ij} = 1$. Therefore $H_i(x,t) = 1 - Q_i(x,t)$ is the survival probability in state i.

In the second term, it collects all cases in which the transition from i to j occurs via another state say $k \neq i$ applying the renewal argument. First, the

probability that the patient stays in state i for a period of length τ and then passes to state k is captured by $Q_{ik}(x,\tau) = w_i(x,\tau)p_{ik}(x,\tau)$. Passing to this new state k can be interpreted as a renewal of the process. Hence the probability that the patient who is in state k at time τ will be in state j at time t is captured by $p_{kj}(\tau,t)$. As the transition from state i to state k could occur any time between 0 and t, all possible transition times τ have been covered by summation over t and the states covered under summation over all states k. The summation in equation (4) must be made over transient states since $p_{kj}(t-\tau)$ is zero whenever k is an absorbing state.

Equation (7) gives the probability that a patient starting in state i will be in j by time t. The probabilities in equation (7) are real quantities of interest in the medical practice and are called the interval transition probabilities of the semi-Markov process. Equation (7) when fitted with data will give the conditional probability of a patient moving from one disease state to another, given the current state. This model will also give the probability of a patient's survival given the current state.

4 HIV patient management model

A large fraction of the economic burden of HIV/AIDS is the medical costs of treating persons with HIV. Medical cost estimates are often based on health care utilization by persons with HIV disease. The costs associated with health care utilization in each disease stage are summed across all disease stages from infection to death.

Let c_i be cost incurred by maintaining an HIV patient in state i, then $c_i(t)$ is the cost incurred in maintaining an HIV patient in state i for a period of time t. Also $c_i(x,t)$ is the cost incurred in maintaining an HIV patient in state i given that the patient enters this state at time s and stays in state i until time t. Let c_{ij} be the cost incurred by treating an HIV patient to move from state i to state j, then $c_{ij}(t)$ is the cost incurred by treating an HIV patient to move from state i to state j at time t. Hence, $c_{ij}(x,t)$ is the cost incurred in treating an HIV patient to move from state i to state j given that the patient entered state i at time s and goes to the state j at time t.

Suppose an HIV patient incurs lump sum treatment cost of d_i while in state i. Let w_i be the mean sojourn time in state i. Then the average treatment cost per unit time of the patient in state i is

$$c_i = \frac{d_i}{w_i} \tag{8}$$

Now let C be the total treatment cost incurred by a patient during their lifetime, then using equations 1, 3 and 8 we have

$$
\begin{aligned}
C(x,t) &= \sum_{j \neq i}^{4} Q_{ji}(x,t)c_{ji}(x,t) + H_i(x,t)\frac{d_i}{w_i}(x,t) \\
&= \sum_{j \neq i}^{4} Q_{ji}(x,t)c_{ji}(x,t) + [1 - Q_i(x,t)]c_i(x,t) \quad (9)
\end{aligned}
$$

where $1 - Q_i(x,t)$ is the probability that the patient stayed in state i during the period of time t and c_{ji} is the cost of moving the patient from worse state to state i.

$C(x,t)$ in equation (9) is the total Lifetime Treatment Cost. This model when applied to real data will give the cost of managing an HIV patient given the starting state, the treatment combination which incurs minimum cost and also the treatment combination which is most effective at each state.

5   HIV patient treatment reward model

Let r_i be the reward gained by an HIV patient for remaining in state i; this reward doesn't change with time and the future transition, $r_i(t)$ be the reward an HIV patient gains for remaining in state i at time t, $r_i(s,t)$ is the benefit gained by an HIV patient for remaining in state i given that the patient enters this state at time s and goes to the state j at time t (non-homogeneity). Let r_{ij} be reward gained by an HIV patient for transition from state i to state j, $r_{ij}(t)$ be the reward an HIV patient gains for transition from state i to state j at time t, $r_{ij}(s,t)$ is the benefit gained by an HIV patient for transition from state i to state j given that the patient entered state i at time s and goes to the state j at time t (non-homogeneity).

Let w_i be the mean sojourn time in state i. Then the average reward of the patient in state i is

$$
r_i = \frac{r_i}{w_i} \quad (10)
$$

In this paper, the HIV treatment reward is regarded as the quality-adjusted lifetime of a patient. Now let R be the total reward gained by an HIV patient through treatment during their lifetime, then using equations 1, 3 and 10 we have

$$
\begin{aligned}
R_i(s,t) &= \sum_{j \neq i}^{4} Q_{ji}(x,t) r_{ji}(x,t) + H_i(x,t) \frac{r_i}{w_i}(x,t) \\
&= \sum_{j \neq i}^{4} Q_{ji}(x,t) r_{ji}(x,t) + [1 - Q_i(x,t)] r_i(x,t) \quad\quad (11)
\end{aligned}
$$

Where $1 - Q_i(t)$ is the probability that the patient stayed in state i during the period of time t.

$R_i(s,t)$ in equation (11) is the reward a patient gains in remaining in state i. This model when applied to real data will give the state, which a patient should be maintained so that they remain healthy, non-infectious and productive to the society. Also the model will show the optimal/effective time to initiate treatment, which can be used to give advice on how to handle the HIV infecteds given their states.

6 Conclusion

In this study, stochastic models were formulated and analyzed for HIV evolution dynamics using Non homogeneous Semi-Markov processes with the aim of evaluating the cost of anti-retroviral therapy (ART) for an HIV infected patient under ART follow-up.

Semi-Markov process was used to determine the conditional probability of an HIV patient moving from one disease state to another. The transition probability model shows the probability that a patient starting in state i will be in state j within the tth year of their life, which when fitted with data will give the conditional probability of a patient moving from one disease state to another, given the current state. This model also gives the probability of a patient's survival. The disease management cost model is formulated, which gives the cost of managing an HIV patient to a certain disease state. The HIV cost model when applied to real data will give the treatment combination which incurs minimum cost and also the treatment combination which is most effective at each state. Treatment reward model was formulated using non-homogeneous semi-Markov reward process (NHSMRP). This reward model when applied to real data will give the state, which a patient should be maintained so that they remain healthy, non-infectious and productive to the society. The model will show the optimal or effective time to initiate treatment, which can be used to give advice on how to handle the HIV infecteds given their states and

also it might uncover new intervention strategies to help prevent or eradicate infection.

The study does not utilize all the possibilities of the semi-Markov process, by means of backward recurrence time process, it is possible to assess different transition probabilities as a function of the duration inside the states. Further more, it is recommended that real data may be used in order to test the efficacy of the models. Moreover, considering the patient's age in formulating semi-Markov model for HIV evolution dynamics can be the object of future research.

References

[1] P K Andersen, S Esbjerg, and T I Sorensen. Multi-state models for bleeding episodes and mortality in liver cirrhosis. *Stat Med*, 19:587–599, 2000.

[2] Vlad Barbu, Michel Boussemart, and Nikolaos Limnios. Discrete-time semi-markov model for reliability and survival analysis. *Communications in Statistics - Theory and Methods*, 33(11):2833–2868, January 2004.

[3] Vlad Barbu and Nikolaos Limnios. Empirical estimation for discrete-time semi-markov processes with applications in reliability. *journal of Nonparametric Statistics*, 18(7-8):483–498, 2006.

[4] CDC. *WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children.* World Health Organization, [Geneva], 2007.

[5] Pai-Lien Chen and Hsiao-Chuan Tien. Semi-markov models for multistate data analysis with periodic observations. *Communications in Statistics: Theory and Methods*, 33(3):475–486, 2004.

[6] V DeGruttola and S W Lagakos. Analysis of doubly-censored survival data with application to AIDS. *Biometrics*, 45:1–11, 1989.

[7] Giuseppe Di Biase, Guglielmo D'Amico, Arturo Di Girolamo, Jacques Janssen, Stefano Iacobelli, and Raimondo Manca. A stochastic model for the HIV/AIDS dynamic evolution. *Mathematical Problems in Engineering, Hindawi Publishing Corporation*, 2007:1–15, 2007.

[8] Gregg E Dinse and S W Lagakos. Athe analysis of partially-censored data from a first-order semi-markov model. *Journal of Statistical Computation and Simulation*, 11:209–222, 1980.

[9] Sylvie Escolano, Jean-Louis Golmard, Anne-Marie Korinek, and Alain Mallet. A multi-state model for evolution of intensive care unit patients: Prediction of nosocomial infections and deaths. *Statistics in Medicine*, 19(24):3465–3482, 2000.

[10] Yohann Foucher, Eve Mathieu, Saint-Pierre Philippe, Jean-Francois Durand, and Jean-Pierre Daur'es. A semi-markov model based on generalized weibull distribution with an illustration for HIV disease. *Biometrical Journal*, **47**(6):825–833, 2005.

[11] Halina Frydman. A nonparametric estimation procedure for a periodically observed three-state markov process, with applications to AIDS. *Journal of the Royal Statistical Society*, 54:853–866, 1992.

[12] Halina Frydman. Semiparametric estimation in a three-state duration-dependent markov model from interval-censored observations with application to AIDS data. *Biometrics*, 51:502–511, 1995a.

[13] Halina Frydman. Nonparametric estimation of a markov 'illness-death' process from interval-censored observations, with application to diabetes survival data. *Biometrika*, 82:773–789, 1995b.

[14] Jacques Janssen and Raimondo Manca. Numerical solution of non-homogeneous semi-markov processes in transient case*. 3(3):271–293, 2001.

[15] Pierre Joly and Daniel Commenges. A penalized likelihood approach for a progressive three-state model with censored and truncated data: Application to AIDS. *Biometrics*, 55:887–890, 1999.

[16] M Kang and S W Lagakos. Evaluating the role of human papillomavirus vaccine in cervical cancer prevention. *Stat Methods Med Res*, 13:139–155, 2004.

[17] M Kang and S W Lagakos. Statistical methods for panel data from a semi-markov process, with application to hpv. *Biometrics*, 8:252–264, 2007.

[18] S W Lagakos, C J Sommer, and M Zelen. Semi-markov models for partially censored data. *Biometrika*, 65:311–318, 1978.

[19] J F Lawless and D Y T Fong. State duration models in clinical and observational studies. *Statistics in Medicine*, 18:2365–2376, 1999.

[20] E Mathieu, Y Foucher, P Dellamonica, and JP Daures. Parametric and non homogeneous semi-markov process for hiv control, methodology and computing. *Applied Probability*, :389–397, 2007.

16

[21] Charles J Mode and Gary T Pickens. Computational methods for renewal theory and semi-markov processes with illustrative examples (c/r: 89v43 p130-131). *The American Statistician*, 42:143–152, 1988.

[22] E H Moore and R Pyke. Estimation of the transition distributions of a markov renewal process. *The Annals of Mathematical Statistics*, 39:411–425, 1968.

[23] G A Satten. Estimating the extent of tracking in interval-censored chain-of-events data. *Biometrics*, 55:1228–1231, 1999.

[24] G A Satten and M R Sternberg. Fitting semi-markov models to interval-censored data with unknown initiation times. *Biometrics*, 55:507–513, 1999a.

[25] M R Sternberg and G A Satten. Discrete-time nonparametric estimation for semi-markov models of chain-of-events data subject to interval censoring and truncation. *Biometrics*, 55:514–522, 1999b.