DE GRUYTER OPEN

DOI: 10.2478/bile-2014-0002

Biometrical Letters
Vol. 51 (2014), No. 1, 13 -36

Survival probabilities for HIV infected patients through semi-Markov processes

Giovanni Masala, Giuseppina Cannas and Marco Micocci

Department of Economics, University of Cagliari,Via S. Ignazio, 17 – 09123 Cagliari, Italy, e-mail: gb.masala@unica.it

SUMMARY

In this paper we apply a parametric semi-Markov process to model the dynamic evolution of HIV-1 infected patients. The seriousness of the infection is rendered by the CD4+ T-lymphocyte counts. For this purpose we introduce the main features of non-homogeneous semi-Markov models. After determining the transition probabilities and the waiting time distributions in each state of the disease, we solve the evolution equations of the process in order to estimate the interval transition probabilities. These quantities appear to be of fundamental importance for clinical predictions. We also estimate the survival probabilities for HIV infected patients and compare them with respect to certain categories, such as gender, age group or type of antiretroviral therapy. Finally we attach a reward structure to the aforementioned semi-Markov processes in order to estimate clinical costs. For this purpose we generate random trajectories from the semi-Markov processes through Monte Carlo simulation. The proposed model is then applied to a large database provided by ISS (Istituto Superiore di Sanità, Rome, Italy), and all the quantities of interest are computed.

Keywords: semi-Markov process, HIV states, waiting time distribution, evolution equation, survival probabilities, Monte Carlo simulation.

1. Introduction

Homogeneous semi-Markov processes (HSMP) were introduced in the 1950s, independently by Levy (1954) and Smith (1955), with the objective of generalizing Markov processes. Ina Markov process environment, the waiting time distribution functions in each state must be exponential, whereas in a semi-Markov process environment these distributions can be of any type. A detailed theoretical analysis of semi-Markov processes can be found in Howard

(1971a, b), and more recently in Janssen and Manca (2006). These processes have been applied successfully in a wide range of scientific fields.

In turn, non-homogeneous semi-Markov processes (NHSMP) were introduced by Iosifescu Manu (1972).

We will deal here with semi-Markov stochastic models applied in a clinical field. These processes turn out to be a very efficient tool for predicting the dynamic evolution of human immunodeficiency virus (HIV) infection and the probability of an infected patient's survival. This approach has the following advantages with respect to traditional epidemiological models (according to Di Biase et al., 2007a):

- we can consider an arbitrary number of states, linked to the seriousness of the infection;
- all transitions between states are allowed;
- we can consider the randomness of the evolution between all states, as well as the stochastic time spent in each state before a transition occurs;
- model parameters are directly estimated from raw data;
- finally, conclusions consist in certain interval transition probabilities obtained by solving the evolution equations of the process.

We recall that stochastic tools have been widely used in biomedical applications.

As concerns applications of Markov models, we can mention the studies carried out by Longini et al. (1991), Fischl et al. (1987), and Tsiatis et al. (1992). Nevertheless, the Markov framework used in those papers leads to excessively rigid conditions. According to Brookmeyer and Gail (1994) in their AIDS epidemiology basic text, Markov assumptions are not faithful.

For instance, Longini et al. (1991) assume a time homogeneous Markov chain in order to describe the transitions between seven states characterized by CD4 + T cell numbers. One of the main assumptions of this study is that the hazard rate of transitions from one state to another is constant.

The semi-Markov framework allows much more flexible modeling of these phenomena. Indeed, the duration of each state of the disease can be considered

as a random variable. Many authors have recently applied semi-Markov models in the field of biomedicine. For examples see Davidov (1999), Davidov and Zelen (2000), and Foucher (doctoral dissertation, 2007). Specific applications to HIV infection have been examined by Lagakos et al. (1978), Sattenand Sternberg (1999), Sternberg andSatten (1999), Sweeting et al. (2005), Di Biase et al. (2007a; 2009), Joly (1999), D'Amico et al. (2011), and Goshu and Dessie (2013). Foucher et al. (2005) consider the impact of covariates and assume that the duration time process follows a generalized Weibull distribution. Goshu and Dessie (2013) consider a homogeneous non-parametric semi-Markov model applied to a cohort of Felege-Hiwot hospital (Ethiopia).

It is also important to note, following Di Biase et al. (2007a, b; 2009), that physicians consider that HIV evolution satisfies the working hypotheses inherent in the semi-Markov theory.

The natural evolution of HIV infection usually starts with a latency phase. This phase can last for several years. The main characteristic of HIV infection is the gradual depletion of a particular class of lymphocytes named CD4+ (also called "helper lymphocytes"). These lymphocytes play an essential part in the body's immune response to infections. The depletion of CD4+ then causes a weakening of the immune response, which leads to opportunistic infections of significant seriousness. Besides, the presence of plasma viremia is linked to a possible worsening of the disease.

The disease evolves through successive stages (see Levy, 1993), which can be defined according to CD4+ lymphocyte count, viral load and constitutional symptoms (see Centres for Disease Control and Prevention, 1992). The final stage of the disease (Jaffe and Lifson, 1988) is represented by full blown AIDS.

Three general classes of AIDS markers are available: immunological, viral and clinical. On this topic see Goedert (1990), and Brookmeyer and Gail (1994).

Immunological markers are based on concentrations of CD4+ T cells and CD8+ T cells, levels of serum beta2-microgobulin and serum neopterin, and anergy to cutaneous tests for delayed hyposensitivity.

Viral markers aim to evidence the presence or absence of detectable p24 antigen and plasma viremia (detection of infectious HIV in fresh plasma).

Finally, clinical markers concern typical symptoms such as weight loss, candidiasis, persistent diarrhea, herpes zoster, fatigue and night sweats, persistent fever, and oral hairy leukoplakia.

It is now well known that HIV is transmitted primarily by sexual contact, syringe sharing among intravenous drug users, and blood and blood products that have not been correctly screened. From an epidemiological point of view, the disease has a worldwide spread. It is currently estimated that total cases of HIV infection number some 39.5 million, with a peak in Sub-Saharan Africa and East Asian countries (for more details on this subject, see UNAIDS/WHO AIDS Epidemic Update December 2006).

In this paper we consider a non-homogeneous parametric semi-Markov model in order to determine the dynamic evolution of HIV infected patients through four observable states linked to CD4 count (plus an absorbing state, the death of the patient). The waiting time distributions are given by generalized Weibull distributions and the parameters are obtained using the maximum likelihood procedure. We then solve numerically the evolution equations to obtain the transition probabilities and consequently the survival probabilities. Our model thus combines the parametric approach carried out by Foucher et al. (2005) and the survival analysis conducted by Di Biase et al. in a non-parametric approach (2007a, b; 2009). Nevertheless, the most original feature of our study consists in the generation of random trajectories from the semi-Markov process which permits the attachment of a reward structure. We are then able to estimate the therapy costs for a patient with respect to the stochastic evolution of the disease.

The paper is organized as follows. In this section 1 we have introduced the problem and have listed the main applications of semi-Markov models reported in the literature and the basic features of HIV diffusion. Section 2 is devoted to theoretical aspects of semi-Markov processes. The numerical application and

a description of the database used are presented in section 3. Conclusions are stated in section 4.

2. Semi-Markov Processes

The dynamic evolution of HIV infection can be modeled as a sequence of different states based on the severity of the infection. According to the National Center for Infectious Diseases, Division of HIV/AIDS (see Centres for Disease Control and Prevention, 1992), we can define the following four categories depending on the CD4+ T-lymphocyte counts per microliter of blood: state I ($CD4 \geq 500$), state II ($350 \leq CD4 < 500$), state III($200 \leq CD4 < 350$), and state IV ($CD4 < 200$). We add an absorbing state D corresponding to the death of the patient. Patients can move through these five states and all possible transitions are allowed (we have 20 possible transitions; see Figure 1).

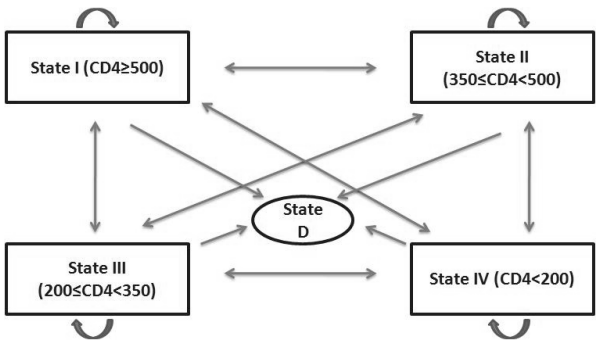

Figure 1. HIV states and transitions

Let us now formalize this problem from a semi-Markov point of view. We first recall the most important aspects of this theory (see Janssen and Manca, 2006).

We denote by $E = \{I, II, III, IV, D\}$ the state space and by $\{\Omega, \mathscr{F}, \mathscr{P}\}$ the usual probability space. Let $m = \text{card}E$ (here $m = 5$). We introduce the following random variables

$$J_n : \Omega \to E \qquad S_n : \Omega \to [0, +\infty) \tag{1}$$

where J_n represents the state at the *n*th transition and S_n represents the chronological time of the *n*th transition. We denote by $N(t)$ the counting process $\{N(t), t \geq 0\}$ which is associated with the point process $(S_n)_{n \in N}$ defined as

$$N(t) = \sup\{n \in N : S_n \leq t\} \tag{2}$$

for every $t \geq 0$.

From this definition, the meaning of the random variable $N(t)$ is clear. It represents the number of transitions occurring in the horizon $[0, t]$.

From the definition of S_n, we can introduce the "duration process" $(X_n)_{n \in N}$ as the following family of random variables:

$$\begin{cases} X_0 = 0 \\ X_{n+1} = S_{n+1} - S_n \end{cases} \tag{3}$$

Consequently, X_{n+1} represents the sojourn time spent in state J_n.

We can now set up the main definition of the theory. The process $(J_n, S_n)_{n \in N}$ is called a "non-homogeneous Markov renewal process" if the following relation holds:

$$\begin{aligned} &P\left(J_{n+1} = j, S_{n+1} \leq t \mid J_n = i, S_n = s, J_{n-1}, S_{n-1}, \ldots, J_0, S_0\right) = \\ &= P\left(J_{n+1} = j, S_{n+1} \leq t \mid J_n = i, S_n = s\right) \end{aligned} \tag{4}$$

Besides, we can associate with the renewal process its non-homogeneous semi-Markov kernel Q (for $j \neq i$) defined as

$$Q = \left[Q_{ij}(s, x)\right] = \left[P\left(J_{N(s)+1} = j, X_{N(s)+1} \leq x \mid J_{N(s)} = i, S_{N(s)} = s\right)\right] \tag{5}$$

We observe that in the second expression in this definition, x represents a duration whereas s represents a chronological time.

It has been proved (Janssen and Manca, 2006) that the probability $p_{ij}(s)$ that a patient performs his or her next transition to state j, given that he or she entered state i at time s, can be computed as:

$$p_{ij}(s) = \lim_{x \to \infty} Q_{ij}(s, x) = P\left(J_{N(s)+1} = j \mid J_{N(s)} = i, S_{N(s)} = s\right) \tag{6}$$

for every $i, j \in E$ and $i \neq j$. Hence we can introduce the transition probability matrix $P(s) = [p_{ij(s)}]_{i,j \in E}$ of the embedded non-homogeneous Markov chain $(J_n)_{n \in N}$. In our application, $P(s)$ is a matrix of order m.

We observe that before entering state j, the patient remains for a certain time x in state i. It is appropriate here to introduce the conditional cumulative distribution function of the waiting time in each state, given the state subsequently occupied. We thus set up:

$$F_{ij}(s,x) = P\left(X_{N(s)+1} \leq x \mid J_{N(s)+1} = j, J_{N(s)} = i, S_{N(s)} = s\right) \tag{7}$$

It can be shown that this function is obtained in the following manner:

$$F_{ij}(s,x) = \begin{cases} Q_{ij}(s,x) / p_{ij}(s) & \text{if} \quad p_{ij}(s) \neq 0 \\ 1 \quad \text{if} \quad p_{ij}(s) = 0 \end{cases} \tag{8}$$

Remark. We recall that a striking difference between a continuous time Markov process and a semi-Markov process concerns the characteristics of the distribution functions $F_{ij}(s,x)$. In a Markov framework, these functions follow a negative exponential distribution. In a semi-Markov framework, the distribution functions $F_{ij}(s,x)$ can be of any type. Consequently, the transition intensity can be decreasing or increasing. It follows that the semi-Markov environment is richer than the Markov environment.

We assume that the distributions of waiting times admit densities, which will be denoted $D(s,x) = \left(d_{ij}(s,x)\right)_{i,j \in E}$. Note that these densities do not depend on t.

We now introduce $H_i(s,x)$, the probability that the process stays in state i for a duration of at least the time x, given that the state i is entered at chronological time s:

$$H_i(s,x) = P\left(X_{N(s)+1} \leq x \mid J_{N(s)} = i, S_{N(s)} = s\right) \tag{9}$$

These probabilities can equivalently be written as

$$H_i(s,x) = \sum_{j=1}^{m} Q_{ij}(s,x) = \sum_{j=1}^{m} p_{ij}(s) \cdot F_{ij}(s,x) \tag{10}$$

We then observe that the marginal cumulative distribution functions of the waiting time in each state depend on both times.

Let us finally introduce the functions $S_i(s,x) = 1 - H_i(s,x)$.

We are now able to introduce the continuous time non-homogeneous semi-Markov process $Z(t)$. This process represents, for each time t, the state occupied by the process, namely (Janssen and Manca, 2006):

$$Z(t) = J_{N(t)} \qquad \text{for } t > 0 \tag{11}$$

We can deduce from these definitions that the semi-Markov process $Z(t)$ is completely determined from knowledge of the transition matrix $\mathbf{P}(t)$ and the duration matrix $\mathbf{D}(s,x)$. In what follows we will see the importance of correctly estimating these two elements.

We observe that two different time scales are involved in these processes. The "chronological time" t refers to an arbitrary origin. In our applications we set $t = 0$ as the first clinical measurement available for the patient. The "internal time" x represents the duration of the time spent in each state.

Remark. The non-homogeneity of the process with respect to time is given by the fact that the jump process $p_{ij}(t)$ depends on the chronological time t. In other words, time has an influence on the transition probabilities.

One of the main objectives of this paper is to define the so-called "interval transition probabilities" given by the following:

$$\begin{aligned} \phi_{ij}(t,x) &= P\left(\text{patient stays in state } j \text{ at time } t+x \mid \text{entered state } i \text{ at time } t\right) = \\ &= P\left(Z(t+x) = j \mid J_{N(t)} = i; S_{N(t)} = t\right) \end{aligned} \tag{12}$$

for every $i, j \in E$.

It can be shown that these probabilities satisfy the obvious property $\phi_{ij}(t,x) \neq \phi_{ij}(t+h, x+h)$ for every $h > 0$.

These probabilities can provide important information for medical applications, so we would like to develop a procedure enabling their determination. It is clear that knowledge of the elements $p_{ij}(t)$ of the transition

matrix and the elements $d_{ij}(s,x)$ of the duration matrix will allow us to determine $\phi_{ij}(t,x)$.

The link between these elements is given by the following fundamental relation, called the evolution equation of the process:

$$\phi_{ij}(t,x) = \delta_{ij} \cdot S_{i.}(t,x) + \sum_{\substack{l=1 \\ l \neq i}}^{m} \int_0^x p_{il}(t) \cdot d_{il}(u) \cdot \phi_{lj}(t+u, x-u)\, du \tag{13}$$

where δ_{ij} is the Kronecker delta.

The terms in the right-hand side of the evolution equation can be interpreted in the following manner.

The first term represents the probability that there is no transition from time t up to $t+x$ given entry of the process into state i at time t. In our medical interpretation, this term gives the probability that the infected patient does not shift to any new stage in a time t (so that his or her clinical situation remains stable).

The second term represents the probability that the system moves into a different state following one of the possible trajectories in some time. In the medical interpretation, this term gives the probability that the infected patient moves to a new clinical situation (the patient's clinical situation improves or worsens).

Let us set $c_{il}(t,x) = p_{il}(t) \cdot d_{il}(x)$. The evolution equation can be rewritten as

$$\phi_{ij}(t,x) = \delta_{ij} \cdot S_{i.}(t,x) + \sum_{\substack{l=1 \\ l \neq i}}^{m} \int_0^x c_{il}(t,u) \cdot \phi_{lj}(t+u, x-u)\, du \tag{14}$$

The following relation also holds:

$$\phi_{ij}(t,0) = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases} \tag{15}$$

In order to solve the evolution equation (13), it is convenient to treat the problem in the discrete framework. Indeed, Corradi et al. (2004) have proved that a numerical solution of the evolution equation can be easily determined

thanks to the quadrature method. They also proved that the numerical solution of the process converges to the discrete time semi-Markov process.

Furthermore, in the same paper, the authors showed that the discrete time process converges to the continuous one if the discretization step tends to zero.

Although time in real-life problems is generally continuous, the real observations of a given system are often made in discrete time, even if the time unit chosen may be infinitesimal. The definitions of semi-Markov processes in the continuous and discrete time cases are very similar. We merely introduce the following probabilities, which are specific to the discrete case (non-homogeneous framework):

$$B = \left[B_{ij}(s,t) \right] = \left[P\left(J_{n+1} = j, S_{n+1} = t \mid J_n = i, S_n = s \right) \right] \tag{16}$$

From the definition of the kernel Q we obtain immediately:

$$b_{ij}(s,t) = \begin{cases} Q_{ij}(s,s) = 0 & \text{if} \quad t = s \\ Q_{ij}(s,t) - Q_{ij}(s,t-1) & \text{if} \quad t > s \end{cases} \tag{17}$$

The evolution equations (13) can be easily discretized (Janssen and Manca, 2006). We obtain the following relations (where h denotes the discretization step):

$$\phi_{ij}(uh,kh) = d_{ij}(uh,kh) + \sum_{l=1}^{m} \sum_{\tau=u+1}^{k} b_{il}(uh,\tau h) \cdot \phi_{lj}(\tau h,kh) \tag{18}$$

The matrix form is the following:

$$\Phi(uh,kh) = D(uh,kh) + \sum_{\tau=u+1}^{k} B(uh,\tau h) \cdot \Phi(\tau h,kh) \tag{19}$$

Janssen and Manca (2006) proved the following two fundamental results: that the equations (19) admit a unique solution; and that the matrices Φ thus obtained are stochastic.

The aforementioned equations can be quickly solved thanks to an iterative procedure we now describe. The following steps are required.

First, let us fix a time horizon $[0,T]$ and a finite subdivision with k intervals of length h. The input data are the transition matrices $P(t)$ and the waiting time distributions $F_{ij}(t)$. The matrices $F(t)$ are $m \times m$ (m denotes the number of states) and they must be evaluated for $t = 1,\ldots,k$ using the distributions previously determined. We make the obvious assumption $F(0) = 0$.

The matrices $Q(t)$ are defined as $Q(t) = F(t) \times P(t)$ ("element by element" product).The matrices $B(t)$ are given by:

$$B(t) = \begin{cases} 0 \quad \text{if} \quad t = 0 \\ Q(t) - Q(t-1) \quad \text{if} \quad t > 0 \end{cases} \tag{20}$$

We then obtain the matrices $H(t)$ in the following way:

$$H_{ij}(t) = \begin{cases} 0 \quad \text{if} \quad i \neq j \\ \sum_{k=1}^{m} Q_{ik}(t) \quad \text{if} \quad i = j \end{cases} \tag{21}$$

These matrices represent the probability of leaving state i within time t.

Finally, we define the matrices $D(t)$ as $D(t) = I - H(t)$.

The full iterative procedure runs as follows:

- We set up the obvious relation $\Phi(0) = I_m$ (identity matrix).
- We deduce $\Phi(h) = B(h) \cdot \Phi(0) + D(h)$.
- Given $\Phi(0), \Phi(h), \ldots, \Phi(kh)$, we deduce recursively

$$\Phi\big((k+1) \cdot h\big) = \sum_{\tau=1}^{k+1} B(\tau \cdot h) \cdot \Phi\big((k+1-\tau) \cdot h\big) + D\big((k+1) \cdot h\big) \tag{22}$$

Finally, we illustrate the Monte Carlo procedure needed to generate sample paths from the semi-Markov process.

We fix the usual time horizon $[0,T]$.The algorithm involves the following steps:

- (1) set $k = 0$ and select the initial state $J_0 = i$ with $i \in \{I, II, III, IV\}$;
- (2) sample the next state from the distribution $p_{J_k,*}$;
- (3) sample the sojourn time $X \sim F_{J_k, J_{k+1}}$;

- (4) set the new jump time $S_{k+1} = S_k + X$;
- (5) if $S_{k+1} \geq T$ or $J_{k+1} = D$ (absorbing state) the cycle ends;
- (6) otherwise set $k = k+1$ and repeat from step (II) until $S_{k+1} \geq T$ or $J_{k+1} = D$.

This algorithm has been revised from Barbu and Limnios (2008).

3. Empirical Application

In this section, we apply the model to a database of HIV infected patients. The database characteristics will be described in subsection 3.1. The chronological time is measured from the first biological measurement. The next subsection 3.2 is devoted to the estimation of the main features of the semi-Markov model, namely the transition probability matrices and the distribution of the waiting times in each state. Subsection 3.3 will then present the numerical solutions of the evolution equations and thus the probabilities of interest for medical analysis. Finally, subsection 3.4 concerns the reward structure as a consequence of the Monte Carlo simulation.

3.1. Database description

Our database contains 2,240 subjects enrolled in the Italian public structures from January 1996 to January 2008,and is provided by ISS (Istituto Superiore di Sanità, Rome, Italy). Our study is based on 25,839 checkups (CD4 counts). This database contains a large amount of information. Each patient is codified and the following data are available: patient identifier, birth date, date of enrolment into the cohort, gender, exposure group (sex between men, hemophiliac, perinatal, etc.), co-infection with hepatitis B or C, vaccination against hepatitis B or C, date of seroconversion, how was seroconversion determined, ART therapy initiation, AIDS diagnosis, country of origin, ethnic/racial group, educational level, date of last clinic visit, date last alive, date of death (primary or contributed causes), date of CD4 count, absolute CD4 count (cells/mm^3), date of HIV RNA measurement, RNA copies (copies/mm^3), antiretroviral drug, date

drug commenced, viral test. The clinical measurements for each subject are scheduled approximately every six months. Some additional measurements are also performed whenever the patient's state worsens.

3.2. Semi-Markov estimation

3.2.1. Estimation of the transition matrices $P(t)$

Consider the time horizon $[0,T]$ with $T=4$ years. We take a partition of this time horizon with monthly step size (total 48 sub-intervals).

The matrices $P(t)$ have already been defined in section 2. Whenever real data is available, these matrices can be estimated in the following way. In our application, we recall that the states are $E=\{I,II,III,IV,D\}$, so that $m=5$ (using the previous notation).

We have to determine a transition matrix $P(t)$ with $t=1/12,2/12,\ldots,4$. The elements $p_{ij}(t)$ are given by

$$p_{ij}(t)=a_{ij}(t)/\sum_{k=1}^{m}a_{ik}(t) \tag{23}$$

where $a_{ij}(t)$ represents the number of direct transitions from state i to state j given entry into state i at time t.

We can also set up a polynomial interpolation in order to approximate the functions $p_{ij}(t)$ with a polynomial (for every $i,j=1,\ldots,m$).

Let us consider an approximation with a polynomial of given order (for example order six, which will be justified later):

$$p_{ij}(t)=\alpha_6\cdot t^6+\alpha_5\cdot t^5+\alpha_4\cdot t^4+\alpha_3\cdot t^3+\alpha_2\cdot t^2+\alpha_1\cdot t+b \tag{24}$$

The coefficients α_i and b are given in Table 1 below (note that state *D* is now denoted by 5). For the sake of brevity, we report only the first transitions.

We observe that the coefficients of monomials of degree five and six are of lower order with respect to the other coefficients. Besides, the Root Mean Square Errors (transition 1-1) for the interpolating polynomials of order three, four, five and six are respectively 0.00278, 0.00235, 0.00234 and 0.00213. We

conclude that the six-order interpolating polynomial is to be preferred. For higher orders, the relative coefficients are notably smaller, so that we finally use a six-order polynomial. This argumentation also concerns the other transitions. This continuous interpolating polynomial then permits the estimation of transition probabilities at every time required by a discretized algorithm.

Table 1. Coefficients of the polynomial regression

Trans.	α_1	α_2	α_3	α_4	α_5	α_6	b
1-1	-0.035631	-0.064013	0.093248	-0.046178	0.010042	-0.000811	0.820992
1-2	0.298632	-0.421427	0.296797	-0.109457	0.020185	-0.001467	0.092354
1-3	-0.070419	0.156754	-0.134130	0.055162	-0.010901	0.000831	0.039994
1-4	-0.043937	0.081442	-0.065998	0.026648	-0.005229	0.000397	0.012771
1-5	-0.148645	0.247244	-0.189916	0.073826	-0.014097	0.001050	0.033890

We show below (Figure 2) the interpolating polynomial for the function $p_{11}(t)$.

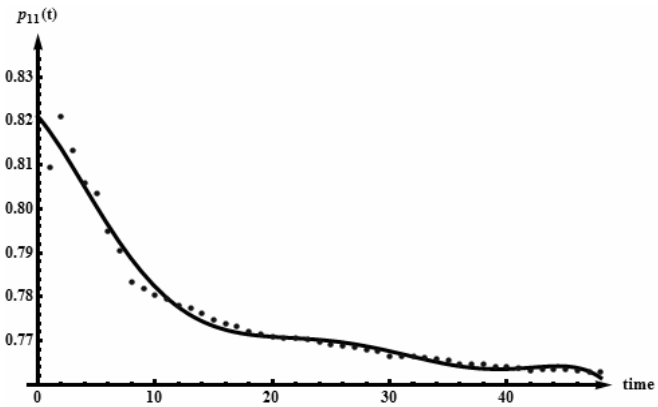

Figure 2. Polynomial regression

3.2.2. The duration time distribution

After determining the number of transitions, we deduce from our database the waiting time in each state (expressed in years). We should then determine a more adequate distribution function of these waiting times. As a typical choice, in accord with Foucher et al. (2005), we consider the generalized Weibull distribution (GW). The authors of that paper assert that this distribution is suitable for this kind of application.

The hazard rate function of the GW is given by:

$$\lambda(x,\nu,\sigma,\theta)=\frac{\nu\cdot\left(\frac{x}{\sigma}\right)^{\nu-1}\cdot\left(\left(\frac{x}{\sigma}\right)^{\nu}+1\right)^{\frac{1}{\theta}-1}}{\theta\cdot\sigma} \tag{25}$$

where ν , σ and θ are real parameters.

The survival function is defined as:

$$S(x,\nu,\sigma,\theta)=e^{1-\left(\left(\frac{x}{\sigma}\right)^{\nu}+1\right)^{1/\theta}} \tag{26}$$

Hence, we deduce the expression of the density function:

$$f(x,\nu,\sigma,\theta)=S(x,\nu,\sigma,\theta)\cdot\lambda(x,\nu,\sigma,\theta) \tag{27}$$

and consequently the distribution function takes the form

$$F(x,\nu,\sigma,\theta)=\int_{-\infty}^{x} f(t,\nu,\sigma,\theta)\,dt \tag{28}$$

We can estimate these three parameters by maximization of the likelihood function

$$L=\prod_{i=1}^{N} f\left(x_i,\nu,\sigma,\theta\right) \tag{29}$$

for each transition, where x_i denotes the time (expressed in years) and N represents the number of observations.

The results are given in Table 2 below (we include the log-likelihood values LogL). For the sake of brevity, we report only the first transitions.

Table 2. GW parameters

Transition	Mean	Variance	ν	σ	θ	Theor. Mean	Theor. Variance	LogL
1-1	0.5293	0.2980	2.8945	0.2164	3.8266	0.5335	0.1775	1,507.50
1-2	0.6269	0.4683	3.4531	0.2214	5.0250	0.6304	0.2897	-641.61
1-3	1.0911	2.4482	2.9607	0.2290	5.8143	1.0793	1.8764	-450.54
1-4	2.7867	7.0690	0.8080	16.0527	0.3282	2.8007	6.9051	-309.02
1-5	2.8532	10.1310	0.7516	4.4461	0.7417	2.8501	11.5468	-136.47

3.3. Numerical results

In this subsection we present the results obtained by solving the evolution equation.

We obtained numerical results taking a time horizon of one year with a monthly step subdivision. The results are given in Table 3.

Table 3. Interval transition probabilities

Prob. (12 months)	I	II	III	IV	D
I	76.63%	17.58%	4.95%	0.51%	0.33%
II	31.85%	44.49%	20.06%	2.92%	0.67%
III	13.33%	24.04%	47.53%	13.25%	1.86%
IV	3.16%	6.75%	20.02%	58.42%	11.66%
D	0.00%	0.00%	0.00%	0.00%	100.00%

For example, let us consider a patient whose initial state is II. After one year, the probability of his state improving is 31.85% while the probability of his state worsening is 22.98%. Besides, the probability of reaching the absorbing state D is 0.67%.

Survival probabilities

Let us define the survival probability for a patient which enters state i at time *zero* as

$$\pi_i(t) = \sum_{j=1}^{m-1} p_{ij}(t) = 1 - p_{iD}(t) \qquad i \in \{I, II, III, IV\} \tag{30}$$

We then perform the following numerical procedure. Consider the time horizon $T = [0,4]$ (years) with a monthly step subdivision. We can represent graphically the values $\pi_i(t)$ for each $i \in \{I, II, III, IV\}$ and $t = 1/12, 2/12, \ldots, 4$.

The result is illustrated graphically in Figure 3 below (in this application, the whole database is involved). The upper line represents $\pi_I(t)$, the second line represents $\pi_{II}(t)$, the third line represents $\pi_{III}(t)$ and finally the lower line indicates $\pi_{IV}(t)$. We deduce immediately that for each value of the time t, the survival probability decreases as the initial state worsens. For each initial state, the survival probability decreases with respect to time t. This decreasing trend is more striking with respect to the worsening of the initial state.

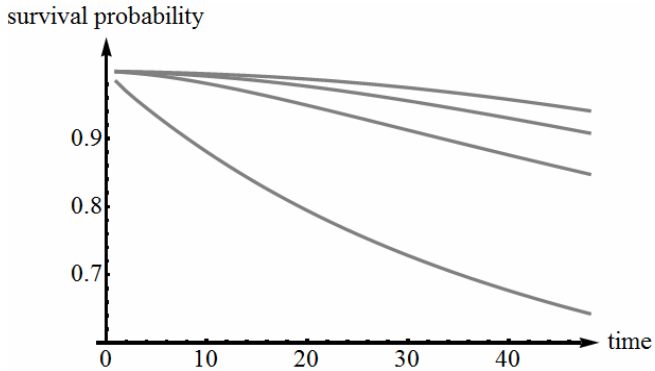

Figure 3. Survival probabilities

Next, we compare the survival probabilities between men and women. We get the following results with respect to the initial state of the patient (see Table 4).

Table 4. Survival probabilities men/women vs. initial state

Time (years)	women (I)	men (I)	women (II)	men (II)	women (III)	men (III)	women (IV)	men (IV)
1	99.52%	99.48%	99.42%	98.85%	98.54%	97.25%	88.07%	86.61%
2	98.66%	98.39%	97.83%	96.81%	95.57%	93.12%	80.74%	76.97%
3	97.30%	96.62%	95.64%	94.07%	92.30%	88.70%	75.34%	69.88%
4	95.50%	94.27%	93.21%	90.92%	89.16%	84.43%	71.17%	64.48%

The same conclusions as stated for Figure 3 also hold here. As a further comment, we note that, for each situation (time and initial state), the survival probability for women is always higher.

These results can be refined by considering different age ranges (as is usual in actuarial practice, for example). We give below the survival probabilities for the ranges 20–30 years and 30–40 years, conditional on the starting state (Tables 5a–5d).

Table 5a. Initial state I

N (years)	M 20-30 Y	W 20-30 Y	M 30-40 Y	W 30-40 Y
1	99.48%	99.52%	99.39%	99.83%
2	98.38%	98.66%	98.13%	98.92%
3	96.62%	97.30%	96.01%	97.31%
4	94.27%	95.50%	93.11%	95.13%

Table 5b. Initial state II

N (years)	M 20-30 Y	W 20-30 Y	M 30-40 Y	W 30-40 Y
1	98.84%	99.42%	99.17%	98.91%
2	96.80%	97.83%	96.99%	96.93%
3	94.05%	95.64%	93.77%	94.31%
4	90.91%	93.21%	89.99%	91.42%

Table 5c. Initial state III

N (years)	M 20-30 Y	W 20-30 Y	M 30-40 Y	W 30-40 Y
1	97.25%	98.54%	97.24%	98.38%
2	93.12%	95.57%	92.53%	95.13%
3	88.69%	92.30%	87.34%	91.56%
4	84.43%	89.16%	82.31%	88.11%

Table 5d. Initial state IV

N (years)	M 20-30 Y	W 20-30 Y	M 30-40 Y	W 30-40 Y
1	86.61%	88.07%	84.69%	89.56%
2	76.97%	80.74%	73.76%	81.35%
3	69.88%	75.34%	65.66%	75.93%
4	64.47%	71.17%	59.48%	72.07%

The interpretation of the results is less straightforward here. When considering initial state IV, we observe that the survival probability of men in the range 20–30 years is higher than for the range 30–40 years. We obtain the opposite conclusion when we consider women in the ranges 20–30 and 30–40 years.

Finally, patients in our database are divided into two categories: patients with antiretroviral therapy and patients without therapy. It is very useful for medical applications to compare the survival probabilities between these two categories. This is indeed a striking indicator that reveals the efficiency of the therapies.

The results are summarized in the following Tables 6a–6d. These contain the estimations with respect to the starting state.

Table 6a. Initial state I

N (years)	therapy	no therapy
1	99.70%	98.88%
2	99.21%	98.03%
3	98.11%	95.45%
4	96.47%	91.79%
5	94.44%	87.74%
6	92.14%	83.04%
7	89.60%	77.06%
8	86.92%	70.19%
9	84.26%	63.84%
10	81.53%	56.64%

Table 6b. Initial state II

N (years)	therapy	no therapy
1	99.35%	97.68%
2	98.30%	96.19%
3	96.40%	91.00%
4	94.00%	84.75%
5	91.44%	79.29%
6	88.85%	73.86%
7	86.16%	67.17%
8	83.39%	60.06%
9	80.73%	54.83%
10	77.97%	47.88%

From these results we can appreciate that the survival probability for patients “with therapy” is higher. We observe a more striking difference for long time horizons.

Table 6c. Initial state III

N (years)	therapy	no therapy
1	98.27%	94.09%
2	95.92%	88.69%
3	92.73%	78.29%
4	89.21%	69.41%
5	86.04%	63.96%
6	83.27%	58.85%
7	80.54%	51.56%
8	77.66%	44.36%
9	75.00%	41.22%
10	72.45%	33.61%

Table 6d. Initial state IV

N (years)	therapy	no therapy
1	90.31%	58.32%
2	84.51%	38.53%
3	77.37%	23.15%
4	72.15%	16.19%
5	69.35%	14.80%
6	67.05%	13.33%
7	63.80%	9.58%
8	60.48%	6.62%
9	58.94%	7.55%
10	55.78%	4.09%

3.4. Reward structure

We can attach a “reward” structure to the process which represents the dynamical evolution of the disease. The goal is to determine the therapy costs for a patient.

In order to perform this estimation, we set up a Monte Carlo procedure to generate sample paths from the semi-Markov process, as described in the previous section.

In order to determine the total cost, we have to assign to each state a specific cost referred to time unity. We can then determine a discounted reward process by considering an interest rate structure. For simplicity, we consider here a constant rate structure given by the instantaneous interest rate δ.

Next we estimate the cost for each simulation, and we then take the expected value over a large number of replications.

Semi-Markov reward processes are also studied in Janssen and Manca (2006). The authors give a recursive formula which permits determination of the cost.

We performed the simulation by assigning the following arbitrary costs for each state (referred to the unit time period, namely one year): 1, 1.5, 2 and 4 for states I, II, III and IV respectively.

The constant interest rate is $\delta = 3\%$, the time horizon is $T = 4$ and the number of simulations is 1,000.

We found the following costs, as given in Table 7 (conditional on the initial state) together with relative confidence intervals at a 95% level.

Table 7. Mean cost

Initial state	Mean cost	95% confidence interval
I	5.11	[5.00,5.22]
II	6.07	[5.93,6.21]
III	7.14	[6.97,7.31]
IV	9.29	[9.05,9.53]

Similar investigations can be carried out considering various additional factors such as patient gender and age group.

Some more references concerning reward structures can be found in Marshall (2007).

4. Conclusions

The HIV model considered in this study clearly concerns a “macroscopic” viewpoint of the dynamical evolution of the disease and it is linked to the CD4 measurement available in our database. This multi-state model consists of four immunological states plus a fifth absorbing state, the death of the patient. We take into consideration all possible transitions. The non-homogeneous semi-Markov model adopted in this paper makes it possible to represent faithfully the dynamical evolution of the disease, and thus we succeed in approximating a real and complex natural process. Besides, we observe that the follow-up time is of fundamental importance for the disease process.

We set a parametric approach through maximum likelihood estimators in order to determine the waiting time distribution functions.

The integral evolution equations of the continuous process are solved in discrete time thanks to an iterative algorithm. This permits us to determine the so-called interval transition probabilities. These probabilities are fundamental in order to perform predictions concerning the clinical evolution of patients.

As a consequence, we can estimate the survival probabilities for HIV infected patients and compare them with respect to certain categories such as gender, age group or type of antiretroviral therapy. This study allows us to deduce certain correlations between survival probabilities and specific factors such as the patient's gender and/or age. As a further improvement, this methodology could be used to compare the efficiency of different antiretroviral therapies.

Finally, a Monte Carlo simulation of semi-Markov processes permits the attachment of a reward structure in order to estimate the therapy costs for a patient.

Acknowledgement

The authors are grateful to Maria Dorrucci, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Roma, Italy, for providing us with the necessary database.

REFERENCES

Barbu V.S., Limnios N. (2008):Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications: Their Use in Reliability and DNA Analysis. Lecture Notes in Statistics N° 191, Springer, New York. DOI: 10.1007/978-0-387-73173-5.

Brookmeyer R., Gail M.H. (1994): AIDS Epidemiology: a Quantitative Approach. Oxford University Press, New York.

Centres for Disease Control and Prevention (1993): Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. MMWR Recommendations and Reports, 41 N° RR-17: 1–19.

Corradi G., Janssen J., Manca R. (2004):Numerical treatment of homogeneous semi-Markov processes in transient case—a straightforward approach. Methodology and Computing in Applied Probability 6: 233–246.

D'Amico G., Di Biase G., Janssen J., Manca R. (2011):HIV Evolution: A Quantification of the Effects Due to Age and to Medical Progress. Informatica 22 (1): 27–42.

Davidov O. (1999): The steady state probabilities for a regenerative semi-Markov processes with application to prevention and screening. Applied Stochastic Models and Data Analysis 15: 55–63.

Davidov O., Zelen M. (2000): Designing cancer prevention trials: a stochastic approach., Statistics in Medicine 19: 1983–1995.

Di Biase G., D'Amico G., Di Girolamo A., Janssen J., Iacobelli S., Tinari N., Manca R. (2007a): Homogeneous semi-Markov model for predicting the HIV disease evolution: a case study. Far Edst. J. Math. Sci. (FJMS) 27: 89–109.

Di Biase G., D'Amico G., Di Girolamo A., Janssen J., Iacobelli S., Tinari N., Manca R. (2007b):A Stochastic Model for the HIV/AIDS Dynamic Evolution. Mathematical problem in Engineering Art. ID 65636, 14 pages. DOI: 10.1155/2007/65636.

Di Biase G., D'Amico G., Janssen J., Manca R. (2009): Patient's Age Depending HIV/AIDS Evolution Analysis by means of a Non Homogeneous Semi-Markov Model. Advances and Applications in Statistics 11: 199–215. ISSN: 0972-3617.

Fischl M.A., Reichmann D.D., Grieco M.H. et al. (1987): The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS related complex. A double blind placebo-controlled trial. New England Journal of Medicine 317: 185–191.

Foucher Y., Mathieu E., Saint-Pierre P., Durand J.F., Daurès J.P. (2005): A semi-Markov model based on generalized Weibull distribution with an illustration for HIV disease. Biometrical Journal 47: 825–833.

Foucher Y. (2007): Modèles semi-markoviens: Application à l'analyse de l'évolution de pathologies chroniques.Doctoral dissertation, Université de Montpellier 1

Goedert J.J. (1990): Prognostic markers for AIDS. Annals of Epidemiology 1: 129–139.

Goshu A.T., Dessie Z.G. (2013): Modelling Progression of HIV/AIDS Disease Stages Using Semi-Markov Processes. Journal of Data Science 11: 269–280.

Howard R.A. (1971a): Dynamic Probabilistic Systems, Markov Models. John Wiley & Sons Vol. 1, New York.

Howard R.A. (1971b): Dynamic Probabilistic Systems, Semi-Markov and Decision Processes. John Wiley & Sons Vol. 2, New York.

Iosifescu Manu A. (1972): Non homogeneous semi-Markov processes, Stud. Lere. Mat. 24: 529–533.

Jaffe H.W., Lifson A.R. (1988): Acquisition and transmission of HIV, Infectious Diseases Clinic of North America 2: 299–306.

Janssen J., Manca R. (2006): Applied Semi-Markov Processes. Springer, New York.

Joly P., Commenges D. (1999): A penalized likelihood approach for a progressive three-state model with censored and truncated data: application to AIDS. Biometrics 55: 887–890.

Lagakos S.W., Sommer C.J., Zelen M. (1978): Semi-Markov models for partially censored Data. Biometrika 65: 311–317.

Levy P. (1954): Processus semi-markoviens. Proceedings of the International Congress of Mathematicians 3: 416–426, Erven P. Noordhoff N.V., Groningen, The Netherlands.

Levy J.A. (1993): Pathogenesis of human immunodeficiency virus infection. Microbiological Reviews 57: 183–289.

Longini I.M., Clark J., Gardner W.S., Brundage J. (1991):The dynamics of CD4+ T lymphocyte decline in HIV infected individuals: A Markov modelling approach. Journal of Acquired Immunodeficiency Syndromes 4: 1141–1147.

Marshall A.H., Shaw B., McClean S.I. (2007): Estimating the costs for a group of geriatric patients using the Coxian phase-type distribution. Statistics in Medicine 26: 2716–2729.

Satten G.A., Sternberg M.R. (1999): Fitting semi-Markov models to interval-censored data with unknown initiation times. Biometrics 55: 507–513.

Smith W.L. (1955): Regenerative stochastic processes. Proceedings of the Royal Society of London Series A. 232: 6–31.

Sternberg M.R., Satten S.A. (1999): Discrete-time nonparametric estimation for semi-Markov models of chain-of-events data subject to interval-censoring and truncation. Biometrics 55: 514–522.

Sweeting M.J., De Angelis D., Aalen O.O. (2005): Bayesian back-calculation using a multi-State model with application to HIV. Statistics in Medicine 24: 3991–4007.

Tsiatis A.A., Dafni U., De Gruttola V. et al. (1992): The relationship of CD4 counts over time to survival of patients with AIDS: Is CD4 a good surrogated marker? Jewell N., Dietz K. and Farewell V (eds.), AIDS Epidemiology: Methodological Issues, Boston, Birkhauser: 257–274.

UNAIDS/WHO AIDS Epidemic Update December 2006 (2006): available at http://www.unaids.org/en/HIV_data/epi2006/default.asp.