

RESEARCH ARTICLE

OPEN ACCESS

A Stochastic Model for the Progression of Chronic Kidney Disease

Noura Anwar*, Mahmoud Riad Mahmoud**

*(ISSR, Cairo University)

**(ISSR, Cairo University)

ABSTRACT

Multistate Markov models are well-established methods for estimating rates of transition between stages of chronic diseases. The objective of this study is to propose a stochastic model that describes the progression process of chronic kidney disease; CKD, estimate the mean time spent in each stage of disease stages that precedes developing end-stage renal failure and to estimate the life expectancy of a CKD patient. Continuous-time Markov Chain is appropriate to model CKD. Explicit expressions of transition probability functions are derived by solving system of forward Kolmogorov differential equations. Besides, the mean sojourn time, the state probability distribution, life expectancy of a CKD patient and expected number of patients in each state of the system are presented in the study. A numerical example is provided. Finally, concluding remarks and discussion are presented.

Keywords-Chronic Kidney Disease , Continuous-Time Markov Chain , Kolmogorov Differential Equations , Expected Time to Absorption, Stochastic Processes .

I. INTRODUCTION

Recently, non-communicable and chronic diseases have become the major causes of morbidity and mortality around the world[1]. One of these diseases is Chronic Kidney Disease which is defined according to the presence or absence of kidney damage and level of kidney function, irrespective of the etiology of kidney disease. Chronic Kidney Disease “CKD” is a worldwide public health problem. It forms a substantial burden for developed societies [21, 11, 22] as well as in developing countries [2, 3, 4, 26, 15, 23]. For example, in Egypt, there are more than 25000 patients with End-Stage Renal Disease “ESRD” and this number have drastically increased over the latest decades [12].

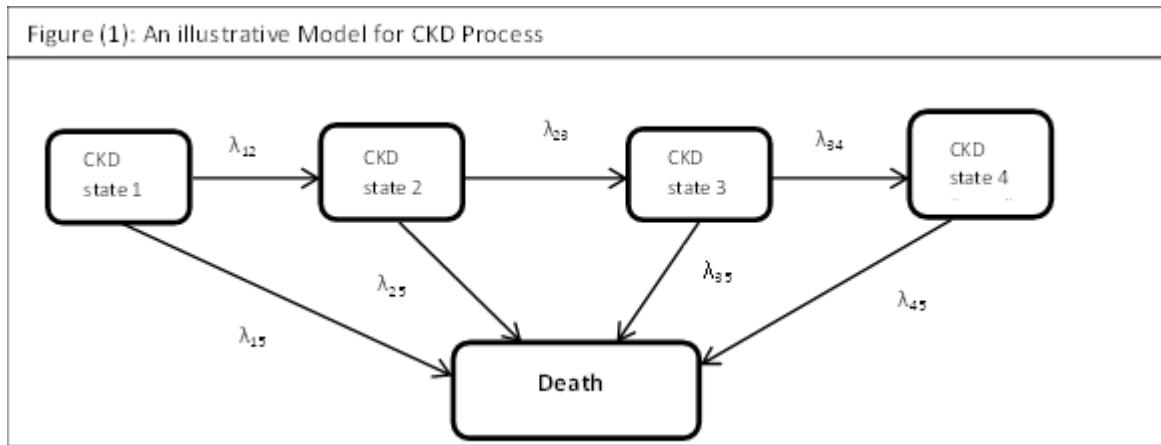
One of the strategies of defeating any chronic disease is to detect it early side by side with the national planning for insuring sufficient treatment of patients. The earlier the CKD is detected the easier to keep the patients in their primary stages and delaying their transition to more severe stages using suitable treatments and suitable lifestyle regime.

Multistate models based on Markov processes are solid methods for estimating rates of transition between stages of diseases. Covariates like age, sex, occupation, previous residence, other chronic diseases, effect of a given intervention ...etc. can be fitted to the transition rates. The expected output of these models help in enhancing the national health policies and forming any preventative strategies of CKD and exploring it in earlier stages where the development of the disease can be revised or prevented. For example, the mean sojourn time that the patient spends in the various states of the process, an important concept of multistate Markov models, may be weighted by cost or utility of a given intervention, then it is used to calculate expected costs and outcomes, thus it allows for comparisons between competing alternatives.

Stochastic models help in understanding the mechanism of diseases in terms of explaining relationships between developing and progressing in disease stages and other relevant covariates. Applications of stochastic processes in medicine and their use in controlling disease-related morbidity and mortality have been attempted by number of authors[13, 18, 14, 5, 27, 25]. The objective of stochastic modeling of diseases vary between research. Number of authors who modeled diseases in order to assess the cost-effectiveness of a new intervention or new technology [6]. Another objective of stochastic models is calculate disease progression and use them in controlling diseases-related mortality. It was used in controlling Cancer-related mortality [18]. Jackson et al.(2003) presented a general Hidden Markov model for simultaneously estimating transition rates and probabilities of stage misclassification when diagnosis of disease stages are subject to error. Later in 2007, Shih and others proposed a method for estimating progression of a chronic disease with multistate properties -Type 2 Diabetes- by unifying the prevalence pool concept with the Markov Process Model. Recently, Begun et al.(2012) proposed a multistate continuous-time nonhomogeneous Markov model for describing patients with

decreased renal function in order to quantify disease progression and its predictor covariates using observable data.

Figure (1): An illustrative Model for CKD Process



The main goal of this study is to propose a stochastic model that describes the progression process of CKD, to estimate the mean time spent in each stage of the disease stages and to estimate the life expectancy of a CKD patient.

II. Methods

Stochastic processes modeling approach is utilized to develop a model of the progression of CKD. General model of disease progression proposed by Jackson et al.(2003). Their general model is consisting of a varying number of transient states and an absorbing state. The model allows for moving progressively from milder to more severe disease stages and vice versa. At the same time it allows moving from any of the disease stage to an absorbing stage.

The general model of disease progression of Jackson et al. (2003) can be used to describe the progression of CKD which is defined, according to the Kidney Disease Outcomes Quality Initiative (KDOQI) classification for CKD, in terms of staged progressive irreversible deterioration of kidney function. CKD process is illustrated graphically in Fig.1.

It is noticed from Fig.1 and from the definition of CKD as a staged progressive irreversible disease that the state space of the progression process is discrete, but the process is continuous with respect to time. The states of CKD that will be considered in the study are defined in Table .1.

Continuous-time Markov chain, "CTMC" is appropriate to model CKD since the patient condition deterioration is continuous in time. A CTMC is said to be homogenous in time if the probability of going from one state to another is independent of the time on which the transition occurs. Homogeneity in time holds true for the process of CKD. Hence, one can assume that the finite homogenous continuous-time Markov chain may be an appropriate model of CKD.

Table (1): Definition of the States of the CKD Model

State No.	State Name	GFR, ml/min
1	Kidney damage with mild reduction in GFR	60 -90
2	Kidney damage with moderate reduction in GFR	30-59
3	Kidney damage with severe reduction in GFR	15-29
4	ESRD implying RRT (regardless of GFR)	< 15
5	Death	

GFR ; glomerular filtration rate: to measure level of kidney function and determine stage of kidney disease.

Depending on the general illustrative model of CKD progression presented in Fig.1, the transition rate matrix of the CKD Progression Model is as follows:

$$V = \begin{pmatrix} -\lambda_{12} - \lambda_{15} & \lambda_{12} & 0 & 0 & \lambda_{15} \\ 0 & -\lambda_{23} - \lambda_{25} & \lambda_{23} & 0 & \lambda_{25} \\ 0 & 0 & -\lambda_{34} - \lambda_{35} & \lambda_{34} & \lambda_{35} \\ 0 & 0 & 0 & -\lambda_{45} & \lambda_{45} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

V is a 5×5 matrix, its elements λ_{ij} are the instantaneous rates of transition from one state to another. It can be noticed that λ_{ij} is independent of time because the CKD process is homogenous with respect to time. Sometimes matrix V is called the generator matrix or the infinitesimal matrix. The elements of V are the model parameters which are population-specific and should be estimated when data is available using the appropriate method of estimation. The initial state probability is given by $\pi_{0i}(t=0) = p(x_0 = S_i)$. This can be written as a vector:

$$\Pi(0) = (\pi_{01} \ \pi_{02} \ \pi_{03} \ \pi_{04} \ 0)$$

This probability defines the probability of being in one of the states of the process at the beginning of the study. Using the analogy of CKD progression process, the initial vector indicates the initial condition of the CKD patients defined as the proportions of patients in each state of the process at the beginning of the study. Entries of initial vector should be nonnegative and their sum should equal to one.

Let us also define the transition probability, $p_{ij}(\tau, t)$, which indicates the probability of being in state i at time τ and would be in state j at time t . The transition probabilities $p_{ij}(\tau, t)$ can also be represented in matrix form.

$$P(\tau, t) = \begin{pmatrix} p_{11}(\tau, t) & p_{12}(\tau, t) & p_{13}(\tau, t) & p_{14}(\tau, t) & p_{15}(\tau, t) \\ p_{21}(\tau, t) & p_{22}(\tau, t) & p_{23}(\tau, t) & p_{24}(\tau, t) & p_{25}(\tau, t) \\ p_{31}(\tau, t) & p_{32}(\tau, t) & p_{33}(\tau, t) & p_{34}(\tau, t) & p_{35}(\tau, t) \\ p_{41}(\tau, t) & p_{42}(\tau, t) & p_{43}(\tau, t) & p_{44}(\tau, t) & p_{45}(\tau, t) \\ p_{51}(\tau, t) & p_{52}(\tau, t) & p_{53}(\tau, t) & p_{54}(\tau, t) & p_{55}(\tau, t) \end{pmatrix}$$

The rows of $P(\tau, t)$ should satisfy the same conditions of the initial state probability vector. The transition probability matrix contains all information necessary to model the movement of a patient among the course of CKD until death.

ACTMC is fully characterized once the transition rate between different states of the system, V , is specified, or conversely when its transition probability matrix is specified along with the expected sojourn time of each state[7].

For simpler CTMC's whose transition rate matrix contains a lot of zero elements, it is appropriate to define transition probability functions in terms of transition intensities through solving system of Kolmogorov differential equations.

Kolmogorov's differential equations play central role in the treatment of Markov processes in continuous time. The forward Kolmogorov differential equations (1) describe the probability distribution of a state in time t keeping the initial point fixed by a so-called "last step analysis". On the other hand, the backward Kolmogorov differential equations (2) describes the transition probabilities in their dependence on the initial point i by the "first step analysis".

$$\frac{\partial}{\partial t} P(\tau, t) = P(\tau, t) \times V \quad (1)$$

with the initial condition

$$P(\tau, \tau) = I$$

where I is the identity matrix. And

$$\frac{\partial}{\partial \tau} P(\tau, t) = -V \times P(\tau, t) \quad (2)$$

with the initial condition

$$P(t, t) = I$$

[8]

Solving system of forward Kolmogorov differential equations yields into the transition probabilities as functions of λ_{ij} 's which are the model parameters. Those parameters are very essential to the solution of the system. These parameters may be estimated by different ways.

Possibility1: When we have a set of observed data where patients are observed in different time points. One can form a likelihood function that reflects all the contributions for all transitions between different observations using the solution of the system of Kolmogorov differential equations as the probability density of transitions

between each pair of states of the system. Then one can obtain maximum likelihood estimates of the model parameters.

Possibility 2: Monte Carlo Markov Chain simulation may be used to simulate life time patient-level trajectories between different states of the system. This requires knowledge of the probability distribution underlying the parameters of the system. This usually requires a lot of time as this application needs a lot of simulations if the data set required is large.

The main objective of this article is to present a solution of the system of forward Kolmogorov equation of CKD process. Then presenting forms of some extractor functions that may be of great importance for the clinicians and health policy makers.

III. Model

3.1 Transition Probability Functions

There are different ways for obtaining an analytical expression for each element of $P(\tau, t)$ in terms of the model parameter V, such as finding the matrix exponentials of the generator matrix V, method of successive approximations, and the spectral methods. For special models, it is possible to calculate an analytic expression for each element of $P(\tau, t)$ by solving the forward Kolmogorov differential equations in (1). This is generally quicker and avoids the possible numerical instability of finding the matrix exponentials [14].

By solving system of equations in (1), we get

$$P_{11}(\tau, t) = e^{-c_1(t-\tau)} \quad (3)$$

where $c_1 = \lambda_{12} + \lambda_{15}$,

$$P_{12}(\tau, t) = \frac{\lambda_{12}}{(c_2 - c_1)} [e^{-c_1(t-\tau)} - e^{-c_2(t-\tau)}] \quad (4)$$

where $c_2 = \lambda_{23} + \lambda_{25}$,

$$\begin{aligned} P_{13}(\tau, t) = & \frac{\lambda_{12} \cdot \lambda_{23}}{(c_2 - c_1)(c_3 - c_1)} e^{-c_1(t-\tau)} - \frac{\lambda_{12} \cdot \lambda_{23}}{(c_2 - c_1)(c_3 - c_2)} e^{-c_2(t-\tau)} \\ & + \frac{\lambda_{12} \cdot \lambda_{23}}{(c_3 - c_1)(c_3 - c_2)} e^{-c_3(t-\tau)} \end{aligned} \quad (5)$$

where $c_3 = \lambda_{34} + \lambda_{35}$,

$$\begin{aligned} P_{14}(\tau, t) = & \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)} e^{-c_1(t-\tau)} - \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)} e^{-c_2(t-\tau)} \\ & + \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34}}{(c_3 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} e^{-c_3(t-\tau)} - \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34}}{(\lambda_{45} - c_1)(\lambda_{45} - c_2)(\lambda_{45} - c_3)} e^{-\lambda_{45}(t-\tau)} \end{aligned} \quad (6)$$

$$\begin{aligned} P_{15}(\tau, t) = & -\frac{a_1}{c_1} e^{-c_1(t-\tau)} - \frac{a_2}{c_2} e^{-c_2(t-\tau)} - \frac{a_3}{c_3} e^{-c_3(t-\tau)} - \frac{a_4}{\lambda_{45}} e^{-\lambda_{45}(t-\tau)} + \\ & \left(\frac{a_1}{c_1} + \frac{a_2}{c_2} + \frac{a_3}{c_3} + \frac{a_4}{\lambda_{45}} \right) \end{aligned} \quad (7)$$

where

$$a_1 = \lambda_{15} + \frac{\lambda_{12} \cdot \lambda_{25}}{(c_2 - c_1)} + \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{35}}{(c_2 - c_1)(c_3 - c_1)} + \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)},$$

$$a_2 = -\frac{\lambda_{12} \cdot \lambda_{25}}{(c_2 - c_1)} - \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{35}}{(c_2 - c_1)(c_3 - c_2)} - \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)},$$

$$a_3 = \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{35}}{(c_3 - c_1)(c_3 - c_2)} + \frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(c_3 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)},$$

$$a_4 = -\frac{\lambda_{12} \cdot \lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(\lambda_{45} - c_1)(\lambda_{45} - c_2)(\lambda_{45} - c_3)},$$

$$P_{21}(\tau, t) = 0 \quad (8)$$

$$P_{22}(\tau, t) = e^{-c_2(t-\tau)} \quad (9)$$

$$P_{23}(\tau, t) = \frac{\lambda_{23}}{(c_3 - c_2)} [e^{-c_2(t-\tau)} - e^{-c_3(t-\tau)}] \quad (10)$$

where $c_3 = \lambda_{34} + \lambda_{35}$,

$$\begin{aligned} P_{24}(\tau, t) = & \frac{\lambda_{23} \cdot \lambda_{34}}{(c_3 - c_2)(\lambda_{45} - c_2)} e^{-c_2(t-\tau)} - \frac{\lambda_{23} \cdot \lambda_{34}}{(c_3 - c_2)(\lambda_{45} - c_3)} e^{-c_3(t-\tau)} \\ & + \frac{\lambda_{23} \cdot \lambda_{34}}{(\lambda_{45} - c_2)(\lambda_{45} - c_3)} e^{-\lambda_{45}(t-\tau)} \end{aligned} \quad (11)$$

$$P_{25}(\tau, t) = -\frac{b_1}{c_2} \cdot e^{-c_2(t-\tau)} - \frac{b_2}{c_3} \cdot e^{-c_3(t-\tau)} - \frac{b_3}{\lambda_{45}} \cdot e^{-\lambda_{45}(t-\tau)} + \left(\frac{b_1}{c_2} + \frac{b_2}{c_3} + \frac{b_3}{\lambda_{45}} \right) \quad (12)$$

where

$$\begin{aligned} b_1 &= \lambda_{25} + \frac{\lambda_{23} \cdot \lambda_{35}}{(c_3 - c_2)} + \frac{\lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(c_3 - c_2)(\lambda_{45} - c_2)}, \\ b_2 &= -\frac{\lambda_{23} \cdot \lambda_{35}}{(c_3 - c_2)} - \frac{\lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(c_3 - c_2)(\lambda_{45} - c_3)}, \\ b_3 &= \frac{\lambda_{23} \cdot \lambda_{34} \cdot \lambda_{45}}{(\lambda_{45} - c_2)(\lambda_{45} - c_3)}, \end{aligned}$$

$$P_{31}(\tau, t) = 0 \quad (13)$$

$$P_{32}(\tau, t) = 0 \quad (14)$$

$$P_{33}(\tau, t) = e^{-c_3(t-\tau)} \quad (15)$$

$$P_{34}(\tau, t) = \frac{\lambda_{34}}{(\lambda_{45} - c_3)} [e^{-c_3(t-\tau)} - e^{-\lambda_{45}(t-\tau)}] \quad (16)$$

$$P_{35}(\tau, t) = -\frac{d_1}{c_3} \cdot e^{-c_3(t-\tau)} - \frac{d_2}{\lambda_{45}} \cdot e^{-\lambda_{45}(t-\tau)} + \left(\frac{d_1}{c_3} + \frac{d_2}{\lambda_{45}} \right) \quad (17)$$

where

$$\begin{aligned} d_1 &= \lambda_{35} + \frac{\lambda_{34} \cdot \lambda_{45}}{(\lambda_{45} - c_3)}, \\ d_2 &= -\frac{\lambda_{34} \cdot \lambda_{45}}{(\lambda_{45} - c_3)}, \end{aligned}$$

$$P_{41}(\tau, t) = 0 \quad (18)$$

$$P_{42}(\tau, t) = 0 \quad (19)$$

$$P_{43}(\tau, t) = 0 \quad (20)$$

$$P_{44}(\tau, t) = e^{-\lambda_{45}(t-\tau)} \quad (21)$$

$$P_{45}(\tau, t) = 1 - e^{-\lambda_{45} (t-\tau)} \quad (22)$$

$$P_{51}(\tau, t) = 0 \quad (23)$$

$$P_{52}(\tau, t) = 0 \quad (24)$$

$$P_{53}(\tau, t) = 0 \quad (25)$$

$$P_{54}(\tau, t) = 0 \quad (26)$$

$$P_{55}(\tau, t) = 1 \quad (27)$$

It makes sense that we have $P_{ij}(\tau, t) = 0$ for $i > j$ as CKD is irreversible progressive disease, and $P_{55}(\tau, t) = 1$ since state "5" represents "Death" which is absorbing.

3.2 Mean sojourn time

Since CKD is one of the progressive diseases, the patient is not supposed to have many visits to a single state. In other words, the mean sojourn time which describes the expected duration of a single stay in a state will be equivalent to the total length of stay or the mean time spent by the patient in a given state of the process.

The mean sojourn time in a state of a CTMC is calculated in terms of transition rates. It is assumed that the sojourn times e_j 's are independent and exponentially distributed random variables with mean $\frac{1}{\lambda_j}$ [9] where $\lambda_j = -\lambda_{jj}$ for $j = 1, \dots, 4$.

Hence, we conclude that

$$e_1 = \frac{1}{\lambda_{12} + \lambda_{15}} \quad (28)$$

$$e_2 = \frac{1}{\lambda_{23} + \lambda_{25}} \quad (29)$$

$$e_3 = \frac{1}{\lambda_{34} + \lambda_{35}} \quad (30)$$

$$e_4 = \frac{1}{\lambda_{45}} \quad (31)$$

3.3 State Probability Distribution

It is important to estimate the state probability distribution of CKD process in order to calculate the probability that the system will be in a particular state at a specific time point t . Let us assume that the state probability distribution, sometimes known by the marginal distribution, of the process at time t is $\Pi(t)$. For homogenous CTMC, $\Pi(t)$ can be evaluated by solving the following system of differential equations

$$\Pi'(t) = \Pi(t) \times V \quad (32)$$

with the initial condition $\Pi(0) = (\pi_{01} \ \pi_{02} \ \pi_{03} \ \pi_{04} \ 0)$.

Generally, the solution of this system of equations depends on the form of V . Our hope for a solution in some special cases depends on V resulting in a simple system of equations.

The solution of (32) are as follows:

$$\pi_1(t) = \pi_{01} e^{-c_1 t} \quad (33)$$

$$\pi_2(t) = \frac{\lambda_{12}\pi_{01}}{(c_2 - c_1)} e^{-c_1 t} + \left[\pi_{02} - \frac{\lambda_{12}\pi_{01}}{(c_2 - c_1)} \right] e^{-c_2 t} \quad (34)$$

$$\begin{aligned} \pi_3(t) &= \frac{\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)} e^{-c_1 t} + \left[\frac{\lambda_{23}\pi_{02}}{(c_3 - c_2)} - \frac{\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)} \right] e^{-c_2 t} \\ &\quad + \left[\pi_{03} - \frac{\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)} - \frac{\lambda_{23}\pi_{02}}{(c_3 - c_2)} + \frac{\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)} \right] e^{-c_3 t} \end{aligned} \quad (35)$$

$$\begin{aligned} \pi_4(t) &= \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)} e^{-c_1 t} \\ &\quad + \left[\frac{\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_2)} - \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)} \right] e^{-c_2 t} \\ &\quad + \left[\frac{\lambda_{34}\pi_{03}}{(\lambda_{45} - c_3)} - \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_3)} - \frac{\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_3)} \right. \\ &\quad \quad \left. + \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} \right] e^{-c_3 t} \\ &\quad + \left[\pi_{04} - \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)} - \frac{\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_2)} \right. \\ &\quad \quad \left. + \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)} - \frac{\lambda_{34}\pi_{03}}{(\lambda_{45} - c_3)} \right. \\ &\quad \quad \left. + \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_3)} + \frac{\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_3)} \right. \\ &\quad \quad \left. - \frac{\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} \right] e^{-\lambda_{45} t} \end{aligned} \quad (36)$$

and

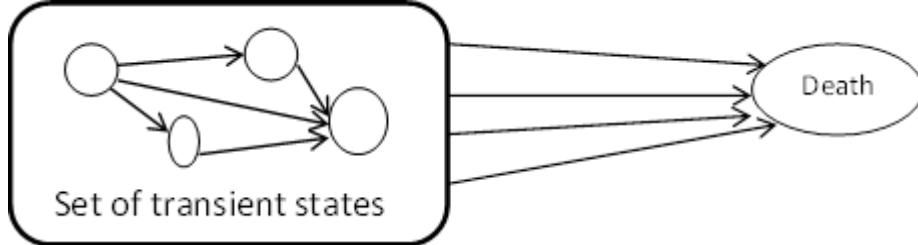
$$\pi_5(t) = \frac{-f_1}{c_1} e^{-c_1 t} + \frac{-f_2}{c_2} e^{-c_2 t} + \frac{-f_3}{c_3} e^{-c_3 t} + \frac{-f_4}{\lambda_{45}} e^{-\lambda_{45} t} + \left(\frac{f_1}{c_1} + \frac{f_2}{c_2} + \frac{f_3}{c_3} + \frac{f_4}{\lambda_{45}} \right) \quad (37)$$

where,

$$\begin{aligned} f_1 &= \lambda_{15}\pi_{01} + \frac{\lambda_{25}\lambda_{12}\pi_{01}}{(c_2 - c_1)} + \frac{\lambda_{35}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)} + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)} \\ f_2 &= \lambda_{25}\pi_{02} - \frac{\lambda_{25}\lambda_{12}\pi_{01}}{(c_2 - c_1)} + \frac{\lambda_{35}\lambda_{23}\pi_{02}}{(c_3 - c_2)} - \frac{\lambda_{35}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)} + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_2)} \\ &\quad - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)} \\ f_3 &= \lambda_{35}\pi_{03} - \frac{\lambda_{35}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)} - \frac{\lambda_{35}\lambda_{23}\pi_{02}}{(c_3 - c_2)} + \frac{\lambda_{35}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)} + \frac{\lambda_{45}\lambda_{34}\pi_{03}}{(\lambda_{45} - c_3)} \\ &\quad - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_3)} - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_3)} + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} \\ f_4 &= \lambda_{45}\pi_{04} - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_1)(\lambda_{45} - c_1)} - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_2)} + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_2)} \\ &\quad - \frac{\lambda_{45}\lambda_{34}\pi_{03}}{(\lambda_{45} - c_3)} \\ &\quad + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} - \frac{\lambda_{45}\lambda_{34}\lambda_{23}\pi_{02}}{(c_3 - c_2)(\lambda_{45} - c_3)} + \frac{\lambda_{45}\lambda_{34}\lambda_{23}\lambda_{12}\pi_{01}}{(c_2 - c_1)(c_3 - c_2)(\lambda_{45} - c_3)} \end{aligned}$$

3.4 life expectancy of a CKD Patient

For CKD process with 5 states, where the first four are transient states and the last is absorbing. We can partition the system as follows:



Also, we can partition the Kolmogorov forward differential equations in (1) as follows:

$$[P'(t)p'_k(t)] = [P(t)p_k(t)] \begin{bmatrix} B & A \\ 0 & 0 \end{bmatrix} \quad (38)$$

Where the matrix B is the transition matrix within transient states, the column vector A is the transition rates from the transient states to the absorbing state. Hence, $A = -B 1^T$, where 1^T is a $(k-1) \times 1$ column vector all its elements are ones.

The system of equations in (38) may be written as follows,

$$\left\{ \begin{array}{l} P'(t) = P(t) B \\ p'_n(t) = P(t) A \end{array} \right. \quad (39-a)$$

$$\left\{ \begin{array}{l} p'_n(t) = P(t) A \\ p_n(t) = P(t) A \end{array} \right. \quad (39-b)$$

Solving (39-a), we get

$$P(t) = P(0)e^{Bt} \quad (40)$$

Substituting by (40) in (39-b), we have

$$p'_n(t) = P(0)e^{Bt} A \quad (41)$$

Note that e^{Bt} is the matrix exponentials of B , defined as follows:

$$e^{Bt} = I + Bt + \frac{1}{2}(Bt)^2 + \frac{1}{3}(Bt)^3 + \dots = \sum_{i=0}^{\infty} \frac{1}{i!} (Bt)^i \quad (42)$$

According to [8] as well as other authors, the solution of the first equation of (39-a) given in equation (40) is an explicit solution of the forward Kolmogorov equation and $P(0)$ is the transition probability matrix at initial time point $t = 0$ which equals to I , the identity matrix.

Given that τ_k is the time to reach the absorbing state from the initial time point, we have

$$F_k(t) = \text{pr}\{\tau_k \leq t\} = \text{pr}\{X(t) = k\} = p_k(t) = 1 - P(t)1^T = 1 - P(0)e^{Bt}1^T \quad (43)$$

Random variables which have cumulative distribution function of such form presented in (43), their mean and other moments can be evaluated using the moment theorem for Laplace transforms.
 First, CTMC with an absorbing state will be presented in Laplace transform such that

$$[sP^*(s) - P(0) \ sp_k^*(s)] = [P^*(s)p_k^*(s)] \begin{bmatrix} B & A \\ 0 & 0 \end{bmatrix} \quad (44)$$

Hence, equation (44) is presented as follows

$$\left\{ \begin{array}{l} sP^*(s) - P(0) = P^*(s) B \\ sp_k^*(s) = P^*(s) A \end{array} \right. \quad (45)$$

And (40) and (41) will be:

$$\left\{ \begin{array}{l} P^*(s) = P(0)(sI - B)^{-1} \\ p_k^*(s) = \frac{1}{s} P^*(s) A = \frac{1}{s} P(0)(sI - B)^{-1} A \end{array} \right. \quad (46)$$

It turns out that:

$$F^*_k(t) = \frac{1}{s} P(0)(sI - B)^{-1} A \quad (47)$$

and

$$f^*_k(t) = s F^*_k(t) = P(0)(sI - B)^{-1} A \quad (48)$$

Then we can evaluate the mean time to absorption (with $A = -B 1^T$),

$$\begin{aligned} E[\tau_k] &= (-1) \left. \frac{df^*_k(t)}{ds} \right|_{s=0} \\ &= (-1) P(0)(sI - B)^{-2} A|_{s=0} \\ &= P(0)(-B)^{-1} 1^T \end{aligned} \quad (49)$$

[7]

Hence, life expectancy of a CKD patient can be evaluated using (49), where

$$B = \begin{pmatrix} -\lambda_{12} - \lambda_{15} & \lambda_{12} & 0 & 0 \\ 0 & -\lambda_{23} - \lambda_{25} & \lambda_{23} & 0 \\ 0 & 0 & -\lambda_{34} - \lambda_{35} & \lambda_{34} \\ 0 & 0 & 0 & -\lambda_{45} \end{pmatrix},$$

and $P(0) = I$.

Thus,

$$\begin{aligned} E(\tau_{15}) &= \frac{1}{\lambda_{12} + \lambda_{15}} + \frac{\lambda_{12}(\lambda_{23} + \lambda_{25})}{\lambda_{12} + \lambda_{15}} + \frac{\lambda_{12}\lambda_{23}(\lambda_{34} + \lambda_{35})}{(\lambda_{12} + \lambda_{15})(\lambda_{23} + \lambda_{25})} \\ &\quad + \frac{\lambda_{12}\lambda_{23}\lambda_{34}(\lambda_{23} + \lambda_{25})(\lambda_{45})}{(\lambda_{12} + \lambda_{15})(\lambda_{34} + \lambda_{35})} \end{aligned} \quad (50)$$

$$E(\tau_{25}) = \frac{1}{\lambda_{23} + \lambda_{25}} + \frac{\lambda_{23}(\lambda_{34} + \lambda_{35})}{\lambda_{23} + \lambda_{25}} + \frac{\lambda_{23}\lambda_{34}\lambda_{45}}{(\lambda_{23} + \lambda_{25})(\lambda_{34} + \lambda_{35})} \quad (51)$$

$$E(\tau_{35}) = \frac{1}{\lambda_{34} + \lambda_{35}} + \frac{\lambda_{34}\lambda_{45}}{\lambda_{34} + \lambda_{35}} \quad (52)$$

$$E(\tau_{45}) = \frac{1}{\lambda_{45}} \quad (53)$$

The mean time to absorption is equivalent to the life expectancy of a CKD patient, therefore $E(\tau_{ij})$, $i = 1, \dots, 4$ can be interpreted as follows: the life expectancy of a patient given that he observed his illness in state i .

3.5 Expected Number of Patients in Each State

let $m(0)$ be the size of patients at initial time point $t = 0$. The initial size of patients $m(0) = \sum_{j=1}^4 m_j(0)$, where $m_j(0)$ is the initial size of patients at state j given that there are $m_5(0) = 0$ patients at state 5 which is "Death" at the initial time point. Assuming that patients move independently within the states of the system and at the end of the time interval $(0, t)$, there are $M_j(t)$ for $j = 1, \dots, 4$ patients in state j at time t and $M_5(t)$ deaths at time t . Depending on Chiang (1968), then the expected number of patients in each state at time t can be computed directly as follows,

$$E\{M_j(t)|m_j(0)\} = \sum_{i=1}^{n-1} m_i(0)p_{ij}(t) \quad \text{for } j = 1, \dots, n-1. \quad (54)$$

and

$$E\{M_5(t)|m_j(0)\} = \sum_{i=1}^{n-1} m_i(0)p_{i5}(t) \quad \forall j \quad (55)$$

Equation (54) can be presented in details as follows

$$E\{M_1(t)|m_1(0)\} = m_1(0)p_{11}(t) \quad (56)$$

$$E\{M_2(t)|m_1(0), m_2(0)\} = m_1(0)p_{12}(t) + m_2(0)p_{22}(t) \quad (57)$$

$$E\{M_3(t)|m_1(0), m_2(0), m_3(0)\} = m_1(0)p_{13}(t) + m_2(0)p_{23}(t) + m_3(0)p_{33}(t) \quad (58)$$

$$\begin{aligned} E\{M_4(t)|m_1(0), m_2(0), m_3(0), m_4(0)\} \\ = m_1(0)p_{14}(t) + m_2(0)p_{24}(t) + m_3(0)p_{34}(t) + m_4(0)p_{44}(t) \end{aligned} \quad (59)$$

and the expected number of deaths will be

$$\begin{aligned} E\{M_5(t)|m_1(0), m_2(0), m_3(0), m_4(0)\} \\ = m_1(0)p_{15}(t) + m_2(0)p_{25}(t) + m_3(0)p_{35}(t) + m_4(0)p_{45}(t) \end{aligned} \quad (60)$$

IV. Numerical Example

This example is adapted from the study of Begun et al. (2013). They used data from a dialysis center serving a region of 310,000 inhabitants. the sample consisted of 2097 CKD patient with at least 2 measurements during January 2005 to December 2010. Our system, showed in figure (1) consists of 5 states. The states of the system are defined in Table.1.

$$\text{Given } V = \begin{pmatrix} -0.16 & 0.15 & 0 & 0 & 0.01 \\ 0 & -0.37 & 0.27 & 0 & 0.1 \\ 0 & 0 & -1.37 & 1.27 & 0.1 \\ 0 & 0 & 0 & -1.8 & 1.8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We find the one-year transition probability matrix as follows

$$P(1) = \begin{pmatrix} 0.85 & 0.11 & 0.02 & 0.003 & 0.02 \\ 0 & 0.69 & 0.19 & 0.06 & 0.06 \\ 0 & 0 & 0.25 & 0.27 & 0.48 \\ 0 & 0 & 0 & 0.16 & 0.84 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The mean time spent by a CKD patient in state 1 approximately equals 6 years and 3 months, while it decreases in state 2to be about 2 years and 9 months. The deterioration in health state of a CKD patient become rapid in more severe stages than mild stages since the mean time spent by a patient in states 3 and 4 of the system is about 8 and 6 months respectively.

Assuming that the initial distribution of patients among the states of the system is $\Pi(0) = (0.4 \ 0.3 \ 0.2 \ 0.1 \ 0)$, then the distribution of this cohort of patients after one year will be approximately as follows: $\Pi(1) = (0.34 \ 0.25 \ 0.11 \ 0.08 \ 0.22)$.

The life expectancy of a patient is about 8 years given that he entered the system in state 1, 5 years for state 2, 2years and 6 months for state 3 and about 7 months for state 4.

Let $m(0) = 1000$ paitients and divided among the five states of the system as follows, $(300 \ 200 \ 400 \ 100 \ 0)$. Then the expected distribution of this cohort of CKD patients after 1 year is $(255 \ 171 \ 144 \ 136 \ 294)$.

V. Discussion

Chronic diseases represent a major concern to health policy makers, especially in developing countries. When a disease is detected at an early stage, it may be more amenable to treatment [13]. Knowledge about the progression of chronic diseases is important because it may help health policy makers to evaluate expected burden of disease in future and to evaluate cost effectiveness of competing interventions. The Markov chain

approach is often used for analyzing progression of diseases by describing the time evolution of an individual in the multistate model [5].

CTMC is more appropriate than discrete-time Markov chain for studying patient progression through successive stages of a chronic disease where transitions may be slow, therefore have small probabilities and cannot be described accurately in discrete time units. Kolmogorov's differential equation plays a key role in defining uniquely CTMC. The solution of Kolmogorov's differential equation depends on the form of the generator matrix of the model. For simpler forms of the generator matrix, the analytical solution of Kolmogorov's forward system of differential equation is achievable. The resulting Mathematical relations between the probability of transition and rate of transition can be used to formulate a likelihood function of transitions, then estimating the elements of the generator function which is the model parameters.

One should take into consideration some important precautions when estimating the model parameters. For example, kinds of data that may exhibit different types of censoring and pay attention to what form of probability density reflects accurately the observed transitions of a patient in order to formulate realistic likelihood of transitions that yields in a maximum likelihood estimate of the generator matrix the closest description to reality of the natural history of the disease. Begun and others, in 2012, considered three kinds of data structure that can be met in such studies and differentiated between the contributions of possible transitions to the full likelihood which is then used to obtain the maximum likelihood estimate to model parameters.

Kalbfleisch and Lawless in 1985, and recently Jackson in 2011, presented a general methods for evaluating the likelihood for general multistate model in continuous time depending on the form of the transition probability matrix. They differentiated between likelihood for intermittently-observed processes, in other words panel data, exactly-observed death times, exactly-observed transition times and censored states.

Some advanced models may be applicable to model natural history of chronic diseases, such as hidden Markov models(see for example, [13]), continuous-time latent Markov model (see for example, [19]) and semi-Markov models (see for example, [10]).

In conclusion, we have presented an explicit form of the transition probability matrix of CKD process with 5 states, the first four of them represent the 2nd, 3rd, 4th , and ESRD of CKD according to the KDOQI classification, and the last state is death. Besides, we presented also explicit forms for some important extractor functions which depends primarily on the transition instantaneous rates.

References

- [1] C. Atkins, *The epidemiology of chronic kidney disease*, Kidney International, 67(94), 2005,S14-S18.
- [2] R. Barsoum, *Renal Replacement Therapy in Egypt*, Saudi Journal for Kidney Diseases and Transplantation, 8 (2), 1997, 152-154.
- [3] R. Barsoum, *Overview: End-Stage Renal Disease in the Developing World*, Artificial Organs, 26 (9), 2002, 737-746.
- [4] R. Barsoum, *End-stage renal disease in North Africa*, Kidney International, 63 (supplement 83),2003, S111-S114.
- [5] Begun, A.Icks, R.Waldeyer, S. Landwehr, M. Koch, G.Giani, *Identification of a Multistate Continuous-Time Nonhomogeneous Markov Chain Model for Patients with Decreased Renal Function*, Medical Decision Making, published online 28 December 2012.
- [6] Briggs, M.Sculpher, *An Introduction to Markov Modelling for Economic Evaluation*, Pharmaeconomics, 13(4),1998, 397-409.
- [7] Cassandras, S.Lafortune, *Introduction to Discrete Event System (2nd edition*, Springer Science and Business Media,2008).
- [8] L. Chiang, *Introduction to Stochastic Processes in Biostatistics(1st edition*, John Willy & Sons, Inc.,1968).
- [9] E.Cinlar, *Introduction to Stochastic Processes*(Prentice-Hall, Inc. Englewood Cliffs, N.J.,1975).
- [10] Y.Foucher, E. Mathieu, P.Dellamonica, J.P.Daures, *Parametric and Nonhomogeneous Semi-Markov Process for HIV Control*, Methodological Computational Applied Probability, 9,2007, 389-397.
- [11] R.A. Hamer, M.ElNahas, *The Burden of Chronic Kidney Disease*,BMJ, 332,2006, 563-567.
- [12] Head of the Registry Committee, *The 9th Annual Report*. The Egyptian Renal Registry, Ain Shams University, Cairo, Egypt, 2008.
- [13] C.H.Jackson, D.L.Sharples,et al. *Multistate Markov Models for disease progression with classification error*, The Statistician, 52(2), 2003, 193-209.
- [14] Jackson, *Multi-state modelling with R: the msm package*, The Journal of Statistical Software, 38(8), 2011, 1-29.
- [15] T.H.Jafar, *The Growing Burden of Chronic Kidney Disease in Pakistan*, The New England Journal of Medicine, 354, 2006, 995-997.

- [16] J.D.Kalbfleisch, J.F. Lawless, *The Analysis of Panel Data under a Markov Assumption*, Journal of American Association, 80(392), 1985, 863-871.
- [17] J.G.Kemeny, J.L. Snell, Finite Markov Chains, (2nd edition, R.R. Donnelley&Sons, Harrisonburg, VA, USA,1976)
- [18] M. Kimmel, O.Y.Gorlova, *Stochastic Models of Progression of Cancer and Their Use in Controlling Cancer-Related Mortality*, International Journal of Applied Mathematics, 13(3), 2003, 279-287.
- [19] J.M. Lange, V.N.Minin, *Fitting and interpreting continuous-time latent Markov models for panel data*, Statistics in Medicine, 32, 2013, 4581-4595.
- [20] A.P. Mark, K.Sameul, *An Introduction to stochastic Modeling*, (4th edition, Elsevier, 2011).
- [21] S. Moeller, S.Gioberge,G. Brown, *ESRD patients in 2001: global overview of patients, treatment modalities and development trends*, Nephrology Dialysis Transplantation, 17(12), 2002, 2071–2076.
- [22] Schieppati, G.Remuzzi, *Chronic renal diseases as a public health problem: Epidemiology, social, and economic implications*, Kidney International, 68(supplement 98),2005, S7–S10.
- [23] H.SharadanBallal, *The Burden of Chronic Kidney disease in A Developing Country*, India, DaVita's Clinical Journal of Nephrologists, 9,2007, 12-16.
- [24] H. Shih, P. Chou, C. Liu,T. Tung, *Estimation of Progression of multi-state Chronic Disease using Markov model and Prevalence Pool Concept*, BMC medical Informatics and Decision Making, 7, 2007,34.
- [25] L. Ventura, G. Carreras, D.Puliti, E.Paci, M. Zappa, G.Miccinesi, *Comparison of multi-state Markov models for cancer progression with different procedures for parameter estimation. An application to breast cancer*, Epidemiology Biostatistics and Public Health,2014, e8729-1 - e8729-9.
- [26] K. Vijay, *End-stage renal disease in developing countries*, Kidney International, 62,2002, 350–362.
- [27] A.M. Yen, H. Chen, *Stochastic Models for multiple Pathways of Temporal Natural History on Co-morbidity of Chronic Diseases*, Computational Statistics and Data Analysis, 57, 2013, 570-588.