Identification of a Multistate Continuous-Time Nonhomogeneous Markov Chain Model for Patients with Decreased Renal Function

Alexander Begun, Andrea Icks, Regina Waldeyer, Sandra Landwehr, Michael Koch and Guido Giani

Med Decis Making 2013 33: 298 originally published online 28 December 2012
DOI: 10.1177/0272989X12466731

Identification of a Multistate Continuous-Time Nonhomogeneous Markov Chain Model for Patients with Decreased Renal Function

Alexander Begun, MSc, Andrea Icks, MD, DPH, MBA, Regina Waldeyer, MSc, Sandra Landwehr, PhD, Michael Koch, MD, Guido Giani, PhD

Objectives. *Markov chain models are frequently used to study the clinical course of chronic diseases. The aim of this article is to adopt statistical methods to describe the time dynamics of chronically ill patients when 2 kinds of data sets—fully and partially observable data are available.* ***Model***. *We propose a 6-state continuous-time Markov chain model for the progression of chronic kidney disease (CKD), where little is known about the transitions between the disease stages. States 1 to 3 of the model correspond to stages III to V of chronic kidney disease in the Kidney Disease Outcomes Quality Initiative (KDOQI) CKD classification. States 4 and 5 relate to dialysis and transplantation (renal replacement therapy), respectively. Death is the (absorbing) state 6.* ***Methods and Data***. *The model can be investigated and identified using Kolmogorov's forward equations and the methods of survival analysis. Age dependency, covariates in the form of the Cox regression, and unobservable risks of transition (frailties) can be included in the model. We applied our model to a data set consisting of all 2097 patients from all renal centers in a region in North Rhine-Westphalia (Germany) in 2005–2010.* ***Results***. *We compared transitions and relative risks to the few data published and found them to be reasonable. For example, patients with diabetes had a significantly higher risk for disease progression compared with patients without diabetes.* ***Conclusions***. *In summary, modeling may help to quantify disease progression and its predictors when only partially observable prospective data are available.* ***Key words:*** *Markov models; mathematical models and decision analysis; survival analysis; statistical methods; database analysis; health service research.* ***(Med Decis Making 2013;33:298–306)***

Knowledge about the progression of chronic disease is important (e.g., to appraise further burden of diseases or to evaluate effectiveness or cost-effectiveness of interventions). The Markov chain approach is often used for analyzing progression of diseases by describing of the time evolution of an individual in the multistate model. This approach is based on the Markovian assumption that the values in any state are only influenced by the values of the state that directly preceded it. A full description of the Markov chain model includes the information about the prevalence of individuals in all states at some initial moment t_0 and either the transition probabilities in the case of the discrete-time model or transition intensities in the case of the continuous-time model. The transition probabilities and intensities, respectively, are usually unknown and can be estimated using statistical concepts such as the maximum likelihood (ML) method or a Bayesian approach. These estimates then can be used to make a projection of the prevalence and incidence over some projection horizon and to compare economic and health outcomes of the strategies of population health intervention.

Chronic kidney disease (CKD) is a significant public health challenge affecting a substantial part of the adult population in developed countries.[1,2] It is characterized by a progressive loss in renal function over

Received 23 March 2012 from Institute of Biometrics and Epidemiology, German Diabetes Center at the Heinrich-Heine-University, Düsseldorf, Germany. Revision accepted for publication 16 September 2012.

Address correspondence to Alexander Begun, Institute of Biometrics and Epidemiology, German Diabetes Center, Auf'm Hennekamp 65, Düsseldorf, D-40225, Germany; telephone: +49(0)2113382401; fax: +49(0)2113382644; e-mail: alexander.begun@ddz.uni-duesseldorf.de.

DOI: 10.1177/0272989X12466731

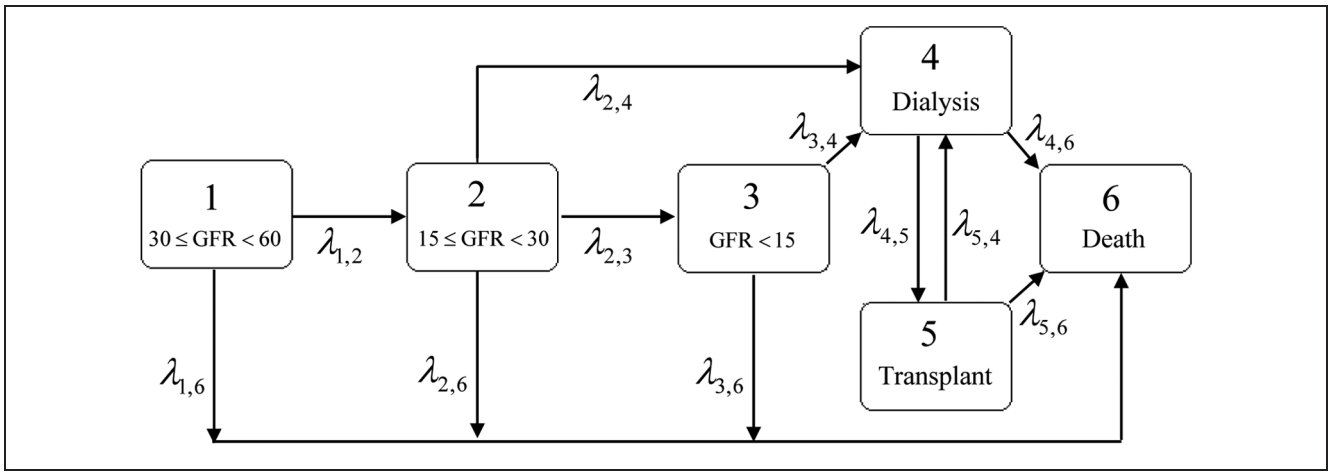

Figure 1 A 6-state Markov model for individuals with chronic kidney disease. GFR, glomerular filtration rate.

months and years. People with high blood pressure, diabetes, and cardiovascular disease have increased risk of CKD.[1–4] The usual way for identifying CKD is to measure the level of serum creatinine. Higher levels of serum creatinine indicate a decreased glomerular filtration rate (GFR), which is a measure for the renal excretion of waste products in urine. There are a number of formulas to calculate GFR values. The recent CKD-EPI (Chronic Kidney Disease–Epidemiology Collaboration) equation may replace the Modification of Diet in Renal Disease (MDRD) formula since it reflects more precisely the GFR with higher values.[5] However, the MDRD formula is most commonly used. In this 4-variable formula, the estimated GFR (eGFR) is a function of serum creatinine, age, sex, and ethnicity.[6,7] It is worth mentioning that after age 40 years, GFR decreases progressively with age in healthy individuals, by about 0.4 to 1.2 mL/min per year.

Following the definition of CKD of the Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines,[8] CKD stages 1 and 2 mean a GFR above 90 and 60 mL/min, respectively, and additional kidney damage, indicated by proteinuria. GFR values below 60 mL/min, 30 mL/min, and 15 mL/min are used to define chronic kidney disease—stages III, IV, and V, respectively. In the end of CKD, in the stage of renal failure, patients need a permanent renal replacement therapy (RRT) such as dialysis or renal transplantation. Data regarding the natural course of CKD are limited, in particular analyzing the progression through the different stages of CKD before renal replacement therapy.[9] Most come from the United States, and it is likely that they cannot be transferred to Europe.[10]

The aim of this article is (1) to adopt statistical methods of Markov chain modeling with different kinds of data and (2) to apply it to a real data set to estimate the course of a population-based cohort of patients passing different stages of CKD. We consider a 6-state continuous-time Markov chain model for the course of CKD with possible nonhomogeneous (age-dependent) transition intensities. The dependence of the transition intensities on age and other observed covariates such as sex or comorbidities is modeled via a Cox-like regression approach. Furthermore, we discuss how the transition intensities can be estimated in case of fully and partially observable data. To estimate the unknown parameters and regression coefficients, we apply the ML method.

METHODS

Markov Model and Data Base

Figure 1 shows the 6-state model. The states have been defined as follows. As long as patients did not receive renal replacement therapy, their state was defined using the eGFR measures (states 1–3). When patients started dialysis or received renal transplant, they went into state 4 or 5. Death was considered as absorbing state 6. The states are described in Table 1.

For our study, we have used the data from a dialysis center covering a region in North Rhine-Westphalia (Germany) with a population of 310,000 inhabitants. All 2097 patients aged 18 years or older (1229 males and 868 females) with diminished GFR (< 60 mL/min) and at least 2 measurements during January

Table 1 Numbering and Definitions of States for a 6-State Model for Individuals with Chronic Renal Disease

State Number	State Name	GFR Values, mL/min
1	Moderate reduction in GFR, stage III, and no RRT	30–59
2	Severe reduction in GFR, stage IV	15–29
3	Established kidney failure, stage V	< 15
4	Dialysis, independent from eGFR	
5	Transplantation (life with transplant), independent from eGFR	
6	Death	

eGFR, estimated glomerular filtration rate; GFR; glomerular filtration rate; RRT, renal replacement therapy (dialysis or renal transplantation).

Table 2 Distribution of the Patients with at Least 2 Estimated Glomerular Filtration Rate (eGFR) Measurements by Sex, Diabetes Status, and Age at First eGFR Measurement

Sex	Diabetes	Nondiabetes	Total
Male			
No. of cases	510	719	1229
Mean (SD) age	69.42 (10.01)	67.49 (13.74)	68.28 (12.36)
Female			
No. of cases	389	479	868
Mean (SD) age	71.39 (11.41)	68.47 (14.35)	69.78 (13.18)
Total			
No. of cases	899	1198	2097
Mean (SD) age	70.27 (10.68)	67.88 (13.99)	68.90 (12.73)

2005 to December 2010 were included. Available variables included the exact data of all eGFR measurements in the study period, date of birth, sex, diabetes status, and start of RRT. The study population is described in Table 2. Possible transitions (or censoring) are given in Table 3.

Death, beginning of dialysis, and renal transplantation are known with exact dates. Hence, transitions $1 \longrightarrow 6$, $2 \longrightarrow 4$, $2 \longrightarrow 6$, $3 \longrightarrow 4$, $3 \longrightarrow 6$, $4 \longrightarrow 5$, $4 \longrightarrow 6$, $5 \longrightarrow 4$, and $5 \longrightarrow 6$ are fully observable. This is not the case for other transitions. Thus, the data are only partially observable.

Analysis of Transition Intensities

Define the age evolution of an individual as a continuous-time Markov process (Y_t), $t \geq 0$, with finite state space $\{1, 2, \ldots, n\}$ for some natural $n \geq 2$. Assume that sample paths of (Y_t) are right-continuous and have left-hand limits. Define an instantaneous rate (hazard) for transition from state i to state j, $j \neq i$, $i,j = 1, \ldots, n$, at age t as

$$\lambda_{i,j}(t) = \lim_{\Delta t \to 0+} P(Y_{(t+\Delta t)-} = j | Y_{t-} = i)/\Delta t.$$

Table 3 Transition Statistics

Transition (or Censoring)	Number
$1 \longrightarrow 1$	842
$1 \longrightarrow 2+$	242
$1 \longrightarrow 3+$	22
$1 \longrightarrow 4$	11
$1 \longrightarrow 6$	19
$2 \longrightarrow 2$	609
$2 \longrightarrow 3+$	159
$2 \longrightarrow 4$	120
$2 \longrightarrow 6$	18
$3 \longrightarrow 3$	286
$3 \longrightarrow 4$	236
$3 \longrightarrow 5$	1
$3 \longrightarrow 6$	11
$4 \longrightarrow 4$	770
$4 \longrightarrow 5$	49
$4 \longrightarrow 6$	422
$5 \longrightarrow 4$	20
$5 \longrightarrow 5$	106
$5 \longrightarrow 6$	9

Destination state coincides with origin state if the continuously observed patient has not changed his or her state in the interval between 2 (not necessarily neighboring) measurements.

Here we assume that the limits exist. If the destination state j is uniquely defined, we have

$$\lambda_i(t) = \sum_{\substack{j=1,\\ j \neq i}}^{n} \lambda_{i,j}(t) = \lim_{\Delta t \to 0+} P(Y_{(t+\Delta t)-} \neq i | Y_{t-} = i)/\Delta t, \quad (1)$$

where $\lambda_i(t)$ is the instantaneous rate for any transition from state i (the total hazard). Notice that the conditional probability of transition $i \to j, j \neq i$, given that transition occurs at age t, is equal to

$$\mu_{i,j}(t) = \lambda_{i,j}(t)/\lambda_i(t).$$

Conditional survival function $S_i(t|t_0)$ is the probability that the process does not leave the state i (an

individual does not move from state i) in the age interval $[t_0, t)$, given that it was in state i at age t_0-. It is easy to check that $S_i(t|t_0)$ is a nonincreasing left-hand continuous function with $S_i(t_0|t_0) = 1$. From (1) and the definition of the survival function, it follows that

$$S_i(t + \Delta t|t_0) = P(Y_\tau = i, \tau \in [t_0, t + \Delta t)|Y_{t_0-} = i) = S_i(t|t_0)[1 - \lambda_i(t)\Delta t + o(\Delta t)].$$

We define $o(\Delta t)$ using the formula $\lim_{\Delta t \to 0+} o(\Delta t)/\Delta t = 0$. Therefore,

$$\lambda_i(t) = -d \ln S_i(t|t_0)/dt = -S_i(t|t_0)^{-1} dS_i(t|t_0)/dt. \quad (2)$$

It means that $S_i(t|t_0)$ has a right-hand derivative and is also a right-continuous function.

From (2) and initial condition $S_i(t_0|t_0) = 1$, it follows that $S_i(t|t_0) = \exp(-H_i(t_0, t))$, where $H_i(t_0, t) = \int_{t_0}^{t} \lambda_i(\tau)d\tau$ is the cumulative hazard. For the (absolutely) continuous conditional random time $T_i(t_0)$ to transition from state i given that the process was in state i at age t_0-, the probability density function corresponding to $S_i(t|t_0)$ is

$$f_i(t|t_0) = -dS_i(t|t_0)/dt = \lambda_i(t)S_i(t|t_0).$$

Denote the probability density function for transition from state i to state $j, j \neq i$, by $f_{i,j}(t|t_0)$, and then

$$f_{i,j}(t|t_0) = \mu_{i,j}(t)f_i(t|t_0) = \lambda_{i,j}(t)S_i(t|t_0).$$

Below we will assume that hazard $\lambda_{i,j}(.)$ depends exponentially on the age t and the observed column vector of measured covariates z, $\lambda_{i,j}(t; z) = a_{i,j} \exp(b_{i,j}t + \beta_{i,j}{}^T z)$. Here, $a_{i,j} > 0$ and $b_{i,j}$ are unknown, $\beta_{i,j}$ is a column vector of unknown regression coefficients, and z stands for explanatory factors (e.g., sex, diabetes status, etc.). This form of dependency corresponds to linear decline in vitality with age for positive $b_{i,j}$[11] and includes a constant intensity as a special case. This formula is usually used for the modeling of mortality, which can be approximated by the Gompertz function in the range of ages 30 to 85 years. This function is also often used for modeling of transition intensities depending on age and covariates. The dependency on the observed covariates is included here in the form of Cox-like regression. In principle, unobserved nonhomogeneity (frailty) can also be included in the model.[12]

In our model, we have $n = 6$ states. Let $p_i(t)$ be the probability of finding an individual at age t in the state i. Denote the 6×6 transition intensity matrix by $\Lambda(t)$ and the 1×6 row vector $(p_1(t), \ldots, p_6(t))$ by $p(t)$. The elements of $\Lambda(t)$ are $\lambda_{i,j}(t)$, and each row sum is equal to zero. In principle, the vector $p(t)$ can be found by solving Kolmogorov's forward equation

$$\frac{dp(t)}{dt} = p(t)\Lambda(t) \quad (3)$$

with initial condition $p(t)|_{t=t_0} = p_0$ for some 1×6 row vector p_0 of probabilities at the initial age t_0 and the transition intensity matrix $\Lambda(t)$:

$$\begin{pmatrix} -\lambda_{1,2}(t) - \lambda_{1,6}(t) & \lambda_{1,2}(t) & 0 & 0 & 0 & \lambda_{1,6}(t) \\ 0 & -\lambda_{2,3}(t) - \lambda_{2,4}(t) - \lambda_{2,6}(t) & \lambda_{2,3}(t) & \lambda_{2,4}(t) & 0 & \lambda_{2,6}(t) \\ 0 & 0 & -\lambda_{3,4}(t) - \lambda_{3,6}(t) & \lambda_{3,4}(t) & 0 & \lambda_{3,6}(t) \\ 0 & 0 & 0 & -\lambda_{4,5}(t) - \lambda_{4,6}(t) & \lambda_{4,5}(t) & \lambda_{4,6}(t) \\ 0 & 0 & 0 & \lambda_{5,4}(t) & -\lambda_{5,4}(t) - \lambda_{5,6}(t) & \lambda_{5,6}(t) \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(both here and below, we use the usual rules for matrix multiplication).

If the matrix $\Lambda(t)$ commutes with its integral $\int_{t_0}^{t} \Lambda(\tau)d\tau$, that is,

$$\Lambda(t)\int_{t_0}^{t} \Lambda(\tau)d\tau = \left(\int_{t_0}^{t} \Lambda(\tau)d\tau\right)\Lambda(t),$$

the solution of the system (3) is given by the formula

$$p(t) = p_0 \exp\left(\int_{t_0}^{t} \Lambda(\tau)d\tau\right).$$

This, for example, is the case when the matrix $\Lambda(t)$ does not depend on t. Another example can be found elsewhere.[13] Notice that for a constant matrix $\Lambda(t) = \Lambda_0$, the matrix $\exp\left(\int_{t_0}^{t} \Lambda_0 d\tau\right)$ is equal to $\sum_{k=0}^{\infty} \frac{(t-t_0)^k}{k!}{\Lambda_0}^k$. If $\Lambda(t)$ depends smoothly on t, the system (3) can be solved numerically for any given accuracy, for example, by Runge-Kutta methods.

We consider 3 basic kinds of data structure that can be met in the study. A discussion of the first two can be found in Welton and Ades.[14]

1. *Fully observable data.* These are the most informative kind of data when we observe individuals continuously and know for an individual the origin and destination states (say, i and $j, j \neq i$) if transition occurs, exact age t_0, and the exact time $t - t_0$ spent by this individual in the origin state before transition occurs. Then, the contribution of this transition to likelihood is equal to $\mu_{i,j}(t)f_i(t|t_0) = \lambda_{i,j}(t)S_i(t|t_0)$. This expression can be interpreted as a probability density of a failure in the survival model with competing risks.[15,16] If $j = i$ (an individual does not leave the state i in the time interval $[t_0, t]$), the contribution to likelihood is equal to $S_i(t|t_0)$. We denote all contributions for fully observable data by $L_l^{(I)}(Data; \omega), l \in D_I$, where D_I is the set of all fully observable transitions and $\omega \in \Omega$ is a vector of unknown parameters.
2. *Partially observable data I.* In this case, we observe an individual at 2 ages: t_0 and t. Assume that the components of the initial vector p_0 are given by δ_{i,i_0}, where $\delta_{i,i_0}$is Kronecker's delta, $\delta_{i,i_0} = 1$ if $i = i_0$ and $\delta_{i,i_0} = 0$if $i \neq i_0$. Then, the solution $p_j(t)$of equation (3) is the probability of finding an individual at age t in state j given that this individual was in state i_0 at age t_0 (note that in this case, $dp_j(t)/dt$ is the probability density function for the transition time from i to j if j is an absorbing state). The contribution to likelihood is equal to $p_j(t)$. In some simple cases, we can calculate $p_j(t)$ without solving equation (3). Consider, for example, an inference of the formula for $p_3(t)$ with initial condition $p(t_0) = \delta_{i,1}$. We observe an individual in the origin state 1 at age t_0 and then in the destination state 3 at age $t > t_0$. Because the destination state 3 can only be reached via state 2, we have $P(Y_t = 3|Y_{t_0} = 1) = \int_{t_0}^{t}\int_{t_0}^{\theta} f_{1,2}(\tau|t_0)f_{2,3}(\theta|\tau)S_3(t|\theta)d\tau d\theta$. If the hazards do not depend on age, we get the formula $P(Y_t = 3|Y_{t_0} = 1) = \frac{\lambda_{1,2}\lambda_{2,3}}{\lambda_2 - \lambda_1}e^{-\lambda_3\Delta t}\left(\frac{e^{(\lambda_3-\lambda_1)\Delta t}-1}{\lambda_3-\lambda_1} - \frac{e^{(\lambda_3-\lambda_2)\Delta t}-1}{\lambda_3-\lambda_2}\right)$, where $\Delta t = t - t_0$. We will enumerate contributions for partially observable data I using notation $L_l^{(II)}(Data; \omega), l \in D_{II}$.
3. *Partially observable data II.* In this case, we know for an individual the origin state i and the destination state $j \neq i$, exact age t_0 in the origin state, and the exact age t at this arrival in the destination state. However, we do not observe the trajectory of this individual over the age interval (t_0, t). Contribution of such transition to likelihood is equal to $\sum_{k \neq j} p_k(t)\lambda_{k,j}(t)$. Here $p_k(t)$ is the probability of finding an individual at age t just before transition in state k given that this individual was in state i at age t_0 (note that $dp_j(t)/dt = \sum_{k \neq j} p_k(t)\lambda_{k,j}(t)$ if j is an absorbing state). Similar to the formula for transition from state 1 to state 3 (see above), we get, for example, that the contribution of the transition $1 \longrightarrow 2 \longrightarrow 4$ to likelihood is equal to $\lambda_{2,4}(t)\int_{t_0}^{t} f_{1,2}(\tau|t_0)S_2(t|\tau)d\tau$ if an individual was in state 1 at age t_0, then arrived in state 4 at age t, and we do not observe age of arriving in state 2. We will enumerate contributions for partially observable data II using notation $L_l^{(III)}(Data; \omega)$, $l \in D_{III}$.

Usually, the 3 kinds of data types mentioned above can be observed in practice. Possible transitions and their contributions to likelihood are given in Table 4. The full likelihood can be calculated by taking the product of all the contributions for all transitions between neighboring measurements

$$L(Data; \omega) = \prod_{l \in D_I} L_l^{(I)}(Data; \omega) \prod_{l \in D_{II}} L_l^{(II)}(Data; \omega) \prod_{l \in D_{III}} L_l^{(III)}(Data; \omega).$$

We get the maximum likelihood estimates of unknown parameters $\hat{\omega}$ from the equation

$$L(Data; \hat{\omega}) = \max_{\omega \in \Omega} L(Data; \omega).$$

In Tables 3 and 4, the sign " + " denotes that we do not know the exact age when a patient arrives in the destination state. For example, for the chain $1 \longrightarrow 2 +$, we know that the patient was in state 1 at age t_0 and in state 2 at age t (partially observable data I).

RESULTS

We have used our model to estimate the age-independent annual transition intensities and the influence on the intensities of 2 explanatory factors: sex and diabetes status. In Table 5, the results for the unknown transition intensities in log scale and the Cox regression parameters are given. They have been derived from stepwise regression with backward elimination based on the likelihood ratio test at level $\alpha = 5\%$. Since the number of patients in states 1, 2, 3, 4, and 5 that either were censored or have not changed the state between 2 (not obligatorily neighboring) measurements is relatively large (see Table 3 for transitions $1 \longrightarrow 1$, $2 \longrightarrow 2$, $3 \longrightarrow 3$, $4 \longrightarrow 4$, and

Table 4 Contribution of Different Transitions to Likelihood ($H_i(t, t+\Delta t) = \int_t^{t+\Delta t} \lambda_i(\tau)d\tau$, $i = 1, \ldots, 5$, $\lambda_1 = \lambda_{1,2} + \lambda_{1,6}$, $\lambda_2 = \lambda_{2,3} + \lambda_{2,4} + \lambda_{2,6}$, $\lambda_3 = \lambda_{3,4} + \lambda_{3,6}$, $\lambda_4 = \lambda_{4,5} + \lambda_{4,6}$, $\lambda_5 = \lambda_{5,4} + \lambda_{5,6}$)

Transition	Formula	Formula (Constant Intensities)
$1 \longrightarrow 2 +$	$\int_t^{t+\Delta t} \lambda_{1,2}(\tau) e^{-H_1(t,\tau)-H_2(\tau,t+\Delta t)} d\tau$	$\lambda_{1,2}(e^{-\lambda_1\Delta t} - e^{-\lambda_2\Delta t})/(\lambda_2 - \lambda_1)$
$1 \longrightarrow 2 \longrightarrow 3 +$	$\int_t^{t+\Delta t} \lambda_{2,3}(\theta) e^{-H_3(\theta,t+\Delta t)} \left(\int_t^{\theta} \lambda_{1,2}(\tau) e^{-H_1(t,\tau)-H_2(\tau,\theta)} d\tau \right) d\theta$	$\frac{\lambda_{1,2}\lambda_{2,3}}{\lambda_2-\lambda_1} e^{-\lambda_3\Delta t} \left(\frac{e^{(\lambda_3-\lambda_1)\Delta t}-1}{\lambda_3-\lambda_1} - \frac{e^{(\lambda_3-\lambda_2)\Delta t}-1}{\lambda_3-\lambda_2} \right)$
$1 \longrightarrow 2 \longrightarrow 4$	$\lambda_{2,4}(t+\Delta t) \int_t^{t+\Delta t} \lambda_{1,2}(\tau) e^{-H_1(t,\tau)-H_2(\tau,t+\Delta t)} d\tau$	$\lambda_{1,2}\lambda_{2,4}(e^{-\lambda_1\Delta t} - e^{-\lambda_2\Delta t})/(\lambda_2 - \lambda_1)$
$1 \longrightarrow 6$	$\lambda_{1,6}(t+\Delta t) \exp(-H_1(t, t+\Delta t))$	$\lambda_{1,6} \exp(-\lambda_1\Delta t)$
$2 \longrightarrow 3 +$	$\int_t^{t+\Delta t} \lambda_{2,3}(\tau) e^{-H_2(t,\tau)-H_3(\tau,t+\Delta t)} d\tau$	$\lambda_{2,3}(e^{-\lambda_2\Delta t} - e^{-\lambda_3\Delta t})/(\lambda_3 - \lambda_2)$
$2 \longrightarrow 4$	$\lambda_{2,4}(t+\Delta t) \exp(-H_2(t, t+\Delta t))$	$\lambda_{2,4} \exp(-\lambda_2\Delta t)$
$2 \longrightarrow 4 \longrightarrow 5$	$\lambda_{4,5}(t+\Delta t) \int_t^{t+\Delta t} \lambda_{2,4}(\tau) e^{-H_2(t,\tau)-H_4(\tau,t+\Delta t)} d\tau$	$\lambda_{2,4}\lambda_{4,5}(e^{-\lambda_2\Delta t} - e^{-\lambda_4\Delta t})/(\lambda_4 - \lambda_2)$
$2 \longrightarrow 6$	$\lambda_{2,6}(t+\Delta t) \exp(-H_2(t, t+\Delta t))$	$\lambda_{2,6} \exp(-\lambda_2\Delta t)$
$3 \longrightarrow 4$	$\lambda_{3,4}(t+\Delta t) \exp(-H_3(t, t+\Delta t))$	$\lambda_{3,4} \exp(-\lambda_3\Delta t)$
$3 \longrightarrow 4 \longrightarrow 5$	$\lambda_{4,5}(t+\Delta t) \int_t^{t+\Delta t} \lambda_{3,4}(\tau) e^{-H_3(t,\tau)-H_4(\tau,t+\Delta t)} d\tau$	$\lambda_{3,4}\lambda_{4,5}(e^{-\lambda_3\Delta t} - e^{-\lambda_4\Delta t})/(\lambda_4 - \lambda_3)$
$3 \longrightarrow 6$	$\lambda_{3,6}(t+\Delta t) \exp(-H_3(t, t+\Delta t))$	$\lambda_{3,6} \exp(-\lambda_3\Delta t)$
$4 \longrightarrow 5$	$\lambda_{4,5}(t+\Delta t) \exp(-H_4(t, t+\Delta t))$	$\lambda_{4,5} \exp(-\lambda_4\Delta t)$
$4 \longrightarrow 6$	$\lambda_{4,6}(t+\Delta t) \exp(-H_4(t, t+\Delta t))$	$\lambda_{4,6} \exp(-\lambda_4\Delta t)$
$5 \longrightarrow 4$	$\lambda_{5,4}(t+\Delta t) \exp(-H_5(t, t+\Delta t))$	$\lambda_{5,4} \exp(-\lambda_5\Delta t)$
$5 \longrightarrow 6$	$\lambda_{5,6}(t+\Delta t) \exp(-H_5(t, t+\Delta t))$	$\lambda_{5,6} \exp(-\lambda_5\Delta t)$

$5 \longrightarrow 5$), the transition intensities can be biased toward zero. Transitions $1 \longrightarrow 6$, $2 \longrightarrow 6$, $2 \longrightarrow 4$, $3 \longrightarrow 6$, $3 \longrightarrow 4$, $4 \longrightarrow 5$, $4 \longrightarrow 6$, $5 \longrightarrow 4$, and $5 \longrightarrow 6$ can be related to fully observable data; transitions $1 \longrightarrow 2 +$, $1 \longrightarrow 2 \longrightarrow 3 +$, and $2 \longrightarrow 3 +$ – to partially observable data I; and transitions $1 \longrightarrow 2 \longrightarrow 4$, $2 \longrightarrow 4 \longrightarrow 5$, and $3 \longrightarrow 4 \longrightarrow 5$ – to partially observable data II. The mean times in years before the patient leaves the states are equal to 6.44 for state 1, 2.70 for state 2, 0.74 for state 3, 5.01 for state 4, and 13.08 for state 5. We have found that women in state 2 have a lower chance of requiring dialysis (relative risk [RR] = 0.50; confidence interval [CI], 0.34–0.67) and of dying (RR = 0.38; CI, 0.09–0.66) compared with men. Diabetes increases the intensity of transition from state 1 to state 2 (RR = 1.41; CI, 1.13–1.68) and the risk of mortality in state 1 (RR = 3.28; CI, 1.63–6.59), decreases the intensity of transition from state 4 to state 5 (RR = 0.46; CI, 0.20–0.73), and increases the risk of mortality for dialysis patients (RR = 1.18; CI, 1.05–1.31). Sex and diabetes status do not significantly influence other transition intensities. The risks of death and renal replacement therapy increase with each stage. Similar results have been reported elsewhere.[9,10]

We contrasted the proposed method with empirical analysis that allows for estimating 1-year transition probabilities P_{ij} as proportions of individuals who started in state i at age $t-1$ and are in state j at age t. If transition intensities do not depend on age, the transition probability matrix P can be calculated from the transition intensity matrix Λ using the formula $P = \exp(\Lambda)$. In the case of no covariates, the results (presented in Table 6) include estimates of P_{ij} and 95% confidence intervals for these estimates. It is evident that the empirical approach suffers from bias in estimation of some transition probabilities.

DISCUSSION

The use of Markov chain models in population studies has a long history. Bartholomew[17] described an approach for estimating the transition matrix for the discrete-time Markov chain using the Dirichlet distribution. Hazen and Pellissier[18] discussed the use of continuous-time Markov chain models based on stochastic trees. The most informative kind of data for identification of the continuous-time Markov chain model must include observations of all state transitions and sojourn times.[19] Welton and Ades[14] proposed an approach for estimating the underlying rate matrix from partially observed data by using Kolmogorov's forward equations. In a recently published article, Yashin and others[20] studied the joint evolution of health and physiological states and their effects on mortality. The parameters of this process and mortality rate were identified from the observed data with 2 mutually dependent continuous and jumping components.

The Markov chain model can be a useful tool for studying multistate populations. The mathematical

Table 5 Estimates of Unknown Parameters in Log Scale and Their Standard Errors

	$\ln a_{1,2}$	$\ln a_{1,6}$	$\ln a_{2,3}$	$\ln a_{2,4}$	$\ln a_{2,6}$	$\ln a_{3,4}$	$\ln a_{3,6}$	$\ln a_{4,5}$	$\ln a_{4,6}$
Estimate	–1.90	–5.12	–1.72	–1.80	–3.71	0.24	–2.83	–3.56	–1.76
SE	0.08	0.37	0.07	0.09	0.22	0.04	0.24	0.15	0.06
	$\ln a_{5,4}$	$\ln a_{5,6}$	$\beta_{2,4}$	$\beta_{2,6}$	$\nu_{1,2}$	$\nu_{1,6}$	$\nu_{4,5}$	$\nu_{4,6}$	
Estimate	–2.94	–3.74	–0.69	–0.98	0.34	1.19	–0.76	0.17	
SE	0.15	0.29	0.16	0.47	0.11	0.44	0.29	0.07	

Here, $e^{\beta_{i,j}}$ is the female/male relative risk and $e^{\nu_{i,j}}$ is the diabetes/nondiabetes relative risk in transition $i \longrightarrow j$.

Table 6 Comparison of Nonzero 1–Year Transition Probabilities: The Proposed v. Empirical Method

	Proposed Method		Empirical Method	
	Estimate	**95% CI**	**Estimate**	**95% CI**
P_{12}	0.137	0.121–0.154	0.163	0.145–0.182
P_{13}	0.009	0.007–0.011	0.015	0.010–0.021
P_{14}	0.013	0.011–0.016	0.007	0.004–0.013
P_{16}	0.014	0.010–0.019	0.013	0.009–0.019
P_{23}	0.081	0.067–0.096	0.165	0.144–0.188
P_{24}	0.164	0.146–0.184	0.125	0.107–0.145
P_{26}	0.033	0.028–0.041	0.019	0.013–0.027
P_{34}	0.626	0.557–0.690	0.714	0.666–0.756
P_{35}	0.009	0.007–0.011	0.003	0.001–0.020
P_{36}	0.108	0.092–0.131	0.033	0.019–0.058
P_{45}	0.019	0.015–0.024	0.022	0.018–0.028
P_{46}	0.167	0.152–0.184	0.190	0.174–0.207
P_{54}	0.046	0.030–0.070	0.057	0.038–0.085
P_{56}	0.028	0.018–0.045	0.026	0.015–0.045

CI, confidence interval.

description of these processes includes Kolmogorov's forward equation. The continuous-time models are more preferable than the discrete-time ones because they allow for transitions that may have a small probability and, therefore, cannot be observed during a small time unit. Both the ML method and the Monte Carlo Markov chain (MCMC) simulations can be used to estimate the unknown parameters of the model. The MCMC method can be used if we have information about the prior distribution of unknown parameters. Usually, this method needs thousands of simulations, and the computation can take a lot of time if the data set is large. Using censored data can lead to parameter estimates that are biased toward zero. In some cases, when the mechanism of censoring is known, the estimates can be corrected. Another problem is the presence in the data set of a large amount of the partially observed data.[14,21,22] This means that an individual can transit to some state not directly but after walking across a number of states. This problem can be solved by using the forward Kolmogorov equations as discussed in Welton and Ades.[14] If the information about walking between 2 not neighboring states for some individual is available and includes a relatively small number of direct transitions between neighboring states, the transition probability can be easily calculated in a closed form. We have calculated such transition probabilities for walking, including 2 direct transitions (see Table 4). Using these formulas, one can substantially reduce the computation time.

Continuous-time Markov chain models are more realistic compared with discrete-time Markov chain models since they allow for state transitions to occur at any time moment. On the other hand, the discrete-time Markov chain models are more popular in medical decision making and simpler to analyze. It is difficult to directly compare results of the parameter

estimation from the discrete-time and continuous-time Markov models. The transition probability matrix for the discrete-time Markov chain usually has more nonzero elements to be estimated than the matrix of transition intensities for the continuous-time Markov chain. It is because we need to include all observed (maybe nondirect) transitions over the regular time interval in the transition probability matrix. This can lead to loss of statistical power. Our analysis shows that use of empirical approach can also lead to biased estimates.

A continuous-time model can be converted into a stochastic equivalent discrete-time model using the uniformization process. This method allows for efficiently calculating an approximation to the transient distribution.[23] Alternatively, an approach similar to that for analyzing transition intensities in the case of the continuous-time Markov chains model and described above can be proposed for analyzing the discrete-time Markov chains.

Unobservable frailty as a measure of general susceptibility to transition can be included in the model to describe transitions in heterogeneous populations. Besides explaining the lack of fit, the frailty component can be useful for modeling dependency in clustered data. The role of frailty in multistate models and the problem of its identifiability have been discussed in a recently published article.[24]

Our results describe the natural progression of CKD in a population-based cohort of CKD patients. Knowledge about CKD progression, including the different stages from population-based prospective studies, is highly limited, particularly in studies that analyze progression of CKD through different stages.[9] Most of them stem from the United States and probably cannot be transferred to other countries. For example, CKD prevalence has been found to be comparable between the US and European countries, but the transition rates to end-stage renal disease or RRT are 2- to 3-fold higher.[9]

With regard to the internal consistency of our results, they appear very reasonable. Mortality risk increases significantly with state number from state 1 to state 4 and then falls drastically for patients with a transplant. The mortality risk for patients with a transplant does not differ significantly from the mortality risk for patients in state 2. Patients with a transplant have significantly lower risk to undergo dialysis in comparison to patients in states 2 and 3.

We also analyzed risks in different subgroups. The higher relative risk of transition from state 2 to state 4 for males can indicate the association of the GFR decline rate with sex. This association has been found in a number of studies.[25] Similarly females compared with males have a lower mortality risk in state 2. Diabetes increases the transition intensities from state 1 to state 2, as well as risks of mortality in states 1 and 4. Diabetes as a risk factor for RRT has been described in a number of studies,[4,26–31] and it is also known that patients with diabetes have a lesser chance of receiving kidney transplantation.

Some limitations need to be acknowledged and addressed regarding the present study. First, a small number of patients can receive preemptive transplants without first passing through a dialysis state[32] (only 1 patient in our sample). This is also true in the Eurotransplant community, where only 2.4% of patients were preemptively transplanted.[33] Second, we have not included the possibility of improvement in the model, as the part of such improvements is relatively small and hardly could have been separated from the positive fluctuations of the eGFR values. For example, in our sample, only 11 (1.6%) patients under dialysis significantly improved their residual renal function and did not need further dialysis treatment. Third, measurement errors in creatinine values can be substantial and can cause false transitions between states. Methods taking into account the measurement imprecision and allowing for a correct state identification should be developed. Fourth, our population may differ from the CKD populations of other renal centers. However, the incidence of renal replacement therapy as well as the distribution of dialysis strategies was well comparable with national data.[34] Fifth, we did not consider both age of CKD onset and the reason of CKD in our model. Both variables may be included in further developments of the model since they are likely to be associated with disease progression.

In conclusion, we were able to adopt statistical methods for the Markov chain model with different kinds of data and apply it to a real data set to estimate a model based on the course of a population-based cohort of patients with CKD through the different CKD stages. We found reasonable results. Modeling may help to quantify disease progression and its predictors when prospective data are lacking. The estimates can be used to predict incidences and prevalence over some time horizon and to compare economic and health outcomes of public health interventions.

REFERENCES

1. Hamer RA, Meguid Ek Nahas A. The burden of chronic kidney disease. BMJ. 2006;332:563–4.

2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiological, social, and economic implications. Kidney Int. 2005;68(Suppl):7–10.

3. Boddana P, Caskey F, Casula A, Ansell D. UK renal registry 11th annual report (December 2008): Chapter 14 UK renal registry and international comparisons. Nephron Clin Pract. 2009;111:269–76.

4. Hsu CY, Iribarren C, McCulloch CE, Darbinian J, Go AS. Risk factors for end-stage renal disease: 25 year follow-up. Arch Intern Med. 2009;169:342–50.

5. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604–12.

6. Levey AS, Bosch JP, Lewis JB, et al. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Dis Study Group. Ann Intern Med. 1999;130:461–70.

7. Hoefield RA, Kalra PA, Baker PG, et al. The use of eGFR and ACR to predict decline in renal function in people with diabetes. Nephrol Dial Transplant. 2011;26(3):887–92.

8. National Kidney Foundation (NKF) Kidney Disease Outcome Quality Initiative (K/DOQI) Advisory Board. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification: Kidney Disease Outcome Quality Initiative. Am J Kidney Dis. 2002;39(Suppl 2):17–31.

9. Gansevoort RT, Matsushita K, van der Velde M, et al; Chronic Kidney Disease Prognosis Consortium. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes: a collaborative meta-analysis of general and high-risk population cohorts. Kidney Int. 2011;80:93–104.

10. Hallan SI, Coresh J, Astor BC, et al. International comparison of the relationship of chronic kidney disease prevalence and ESRD risk. J Am Soc Nephrol. 2006;17:2275–84.

11. Strehler BL, Mildvan AS. General theory of mortality and aging. Science. 1960;132:14–21.

12. Vaupel JW, Manton KG, Stallard S. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography. 1979;16:439–54.

13. Martin JFP. Some result on matrices which commutes with their derivatives. SIAM J Appl Math. 1967;15(5):1171–83.

14. Welton NJ, Ades AE. Estimation of Markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. Med Decis Making. 2005;25:633–45.

15. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. New York: John Wiley; 1980.

16. Cox DR, Oakes D. Analysis of Survival Data. London: Chapman & Hall; 1984.

17. Bartholomew DJ. Errors of prediction for Markov chain models. J R Stat Soc. 1975;37(3):444–56.

18. Hazen GB, Pellissier JM. Recursive utility for stochastic trees. Operations Res. 1996;44:788–809.

19. Parmigiani G. Modeling in Medical Decision Making: A Bayesian Approach. New York: John Wiley; 2002.

20. Yashin AI, Akushevich I, Arbeev KG, Kulminski A, Ukraintseva S. Joint analysis of health histories, physiological states and survival. Math Popul Stud. 2011;18(4):207–33.

21. Borg S, Persson U, Jess T, et al. A maximum likelihood estimator of a Markov model for disease activity in Crohn's disease and ulcerative colitis for annually aggregated partial observations. Med Decis Making. 2009;30:132–42.

22. Van den Hout A, Matthews FE. Estimating dementia-free life expectancy for Parkinson's patients using Bayesian inference and microsimulation. Biostatistics. 2009;10:729–43.

23. Reibman A, Trivedi K. Numerical transient analysis of Markov models. Comput Operations Res. 1988;15(1):19–36.

24. Putter H, van Houwelingen HC. Frailties in multistate models: are they identifiable? Do we need them? Stat Methods Med Res. 2011 Nov 23. [Epub ahead of print]

25. Eriksen BO, Ingebretsen OC. The progression of chronic kidney disease: a 10-year population-based study of the effects of gender and age. Kidney Int. 2006;69:375–82.

26. Bash LD, Astor BC, Coresh J. Risk of incident ESRD: a comprehensive look at cardiovascular risk factors and 17 years of follow-up in the ARIC study. Am J Kidney Dis. 2010;55:31–41.

27. Dirks JH, De Zeeuw D, Agarwal SK, et al. Prevention of chronic kidney and vascular disease: toward global health equity—the Bellagio 2004 Declaration. Kidney Int. 2005;68(Suppl):1–6.

28. Hoffmann F, Haastert B, Koch M, Giani G, Glaeske G, Icks A. The effect of diabetes on incidence and mortality in end-stage renal disease in Germany. Nephrol Dial Transplant. 2011;26(5):1634–40.

29. Icks A, Haastert B, Genz J, et al. Incidence of renal replacement therapy (RRT) in the diabetic compared to the non-diabetic population in a German region, 2002–2008. Nephrol Dial Transplant. 2011;26(1):264–9.

30. Icks A, Haastert B, Genz J, et al. Time-dependent impact of diabetes on the mortality in patients on renal replacement therapy: a population-based study in Germany, 2002–2009. Diabetes Res Clin Pract. 2011;92(3):380–5.

31. Muntner P, Coresh J, Powe NR, Klag MJ. The contribution of increased diabetes prevalence and improved myocardial infarction and stroke survival to the increase in treated end-stage renal disease. J Am Soc Nephrol. 2003;14:1568–77.

32. Vanrenterghem Y, Jones E. Report based on the centre questionnaire. Nephrol Dial Transplant. 1995;11:28–32.

33. Oosterlee A, Rahmel A. Eurotransplant International Foundation: Annual Report 2011. Leiden, the Netherlands: Eurotransplant International Foundation; 2011.

34. Icks A, Genz J, Giani G, Hoffmann F, Trapp R, Koch M. Incidence of renal replacement therapy (RRT) in the diabetic compared to the non-diabetic population in a German region, 2002–2008. Nephrol Dial Transplant. 2011;26(1):264–9.