Check for updates

Received: 20 April 2022 | Revised: 6 February 2023 | Accepted: 27 February 2023

DOI: 10.1002/sim.9707

RESEARCH ARTICLE

Statistics in Medicine WILEY

Non-homogeneous continuous-time Markov chain with covariates: Applications to ambulatory hypertension monitoring

Joonha Chang | **Hei Kit Chan** | **Jeffrey Lin** | **Wenyaw Chan**

Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston, Houston, Texas, USA

Correspondence
Wenyaw Chan, Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston, Houston, TX, USA.
Email: wenyaw.chan@uth.tmc.edu

Hypertension significantly increases the risk for many health conditions including heart disease and stroke. Hypertensive patients often have continuous measurements of their blood pressure to better understand how it fluctuates over the day. The continuous-time Markov chain (CTMC) is commonly used to study repeated measurements with categorical outcomes. However, the standard CTMC may be restrictive, because the rates of transitions between states are assumed to be constant through time, while the transition rates for describing the dynamics of hypertension are likely to be changing over time. In addition, the applications of CTMC rarely account for the effects of other covariates on state transitions. In this article, we considered a non-homogeneous continuous-time Markov chain with two states to analyze changes in hypertension while accounting for multiple covariates. The explicit formulas for the transition probability matrix as well as the corresponding likelihood function were derived. In addition, we proposed a maximum likelihood estimation algorithm for estimating the parameters in the time-dependent rate function. Lastly, the model performance was demonstrated through both a simulation study and application to ambulatory blood pressure data.

KEYWORDS

hypertension, longitudinal categorical data, non-homogeneous Poisson process, stochastic modeling, time-dependent rate

1 | INTRODUCTION

Hypertension is a major risk factor for numerous health problems including heart disease, heart attack, and stroke. In the United States, high blood pressure affects nearly 78 million people and greatly increases the risk of developing cardiovascular disease.[1-3] Among adults aged 18 and over, 48.1% are expected to experience hypertension,[4] with the prevalence of hypertensive heart failure continuing to increase globally.[5,6] Studies have found correlation between hypertension and other factors such as diet, gender, age, and BMI.[7-10] Some previous studies have modeled hypertension through linear models or machine learning techniques.[11-14] However, linear models are limited in modeling complex data which does not adhere to a parametric form. On the other hand, it is more difficult to draw inference through machine learning models. In general, the transitions between hypertensive states were not frequently examined in the longitudinal setting[15,16]

 © 2023 John Wiley & Sons Ltd.

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Hsm/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In addition, the longitudinal models that are capable of distinguishing statistically significant covariates are not robust enough to predict patient's next outcome. To this end, we propose a modified version of the continuous-time Markov chain (CTMC) to make robust predictions while drawing interpretable inference on the model covariates. Continuous-time Markov chains are probabilistic models that describe a sequence of possible events or states following the Markovian property, where the distribution of the upcoming process depends only on the current state and not on any prior states. It is characterized by the set of possible states and the transition rates or probabilities between each state, often denoted conveniently in matrix form. Application of modeling through Markov chains are found in many different areas of study, especially with a wide variety of phenomena in medical research. They are useful in studying the progression of chronic diseases such as HIV infection, breast cancer, and asthma.[17-20] For example, researchers presented applications of Markov models in cancer screening evaluations by using a three-state model to jointly estimate the screening procedure and the average duration of the pre-clinical phase.[19] In addition, a five-state model using survival analysis techniques was applied to predict the mortality reduction from the screening.[21] Further examples of applications on the Markov chain model include its use in health promotion for nutritional interventions[22] and assessing the natural history of colorectal cancer.[23] Generalized estimation equations (GEE) and generalized linear mixed model (GLMM) are two common non-transitional approaches in analyzing longitudinal discrete outcomes such as binary or count outcomes.[24] While the two approaches provide intuitive interpretation, they often lack the ability to account for transitions between outcomes over time on an individual level. On the other hand, Markovian models account for an individual's prior state in making a prediction for their future outcome. Although different in interpretation, Markovian models are also able to determine the effects of covariates similar to GEE and other types of longitudinal regression models. Through existing computation software packages, the likelihood can be derived to include multiple covariates, which allows for a practical interpretation of the Markov process.[25,26] For example, Ma et al demonstrated that the likelihood can be numerically evaluated using ordinary differential equation solvers to accommodate either Bayesian or maximum likelihood approaches.[27] However, due to the complexity of the likelihood function for general multi-state recurrent continuous-time Markov chains, the exact solutions for the observable target state are limited to simple Markov models. Meanwhile, due to its simplicity, CTMC models are limited to only predicting state outcomes, and the accountability for covariates is strictly restricted.

Most applications of the CTMC in analyzing longitudinal discrete outcomes are limited to stationary models. Several studies used Markov models with stationary transition rates to investigate transition patterns in blood pressure states.[28-30] However, the stationary rate assumption of the CTMC may not hold true for many applications where the actual transition rate changes over time, including the rate of change for blood pressure.[31,32] In addition, several covariates are known to affect the status of hypertension and potentially modify the rate at which blood pressure increases or decreases. However, previous studies have incorporated covariates indirectly, and the analyses are limited; Wu et al provided comparative analysis on covariates based on different Markov chains of gender groups and time frames, but the analysis is indirect and restricted from identifying the direct impact and significance on the hypertension transition rates.[30] A few works have focused on deriving theoretical properties of the non-homogeneous CTMC such as ergodicity, irreducibility, and numerical approximation of solutions and probabilities, but the literature regarding the non-homogeneous CTMC and its application to medical research is limited.[33-36] In addition, there have been no prior studies which derive an exact form of the transition probabilities for the non-homogeneous model. Chen and Zhou[37] attempted modeling the transformation of Alzheimer states and death in non-homogeneous Markov process, but the methodology mostly relied on simulations and the expectation-maximization (EM) algorithm rather than exact derivations.

In this article, we proposed a non-homogeneous continuous-time Markov chain (NH-CTMC) models with multiple covariates. The model specified a time-dependent transition rate which may be more suitable when the stationary rate assumption of the CTMC is not met. We algebraically derived the transition probability matrix and the corresponding likelihood function, and presented a maximum-likelihood based algorithm for estimating the parameters of the time-dependent rate function. The model was applied to ambulatory blood pressure data from the Dietary Approaches to Stop Hypertension (DASH) study. In addition, effect of the covariates on the transition of hypertension states were examined. The article is presented as follows: we began with a brief introduction to the Markov chain specific to our setting. We then derived the probability transition functions and the corresponding likelihood functions. Lastly, we described the algorithm used to simulate our NH-CTMC model, verified our model through the simulation results, and applied the proposed model to the data associated with DASH study.

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2 | METHODS

In this section, we propose the probability transition matrices based on the following setup. Consider a two-state NH-CTMC with the infinitesimal matrix $\mathbf{Q}(t)$:

$$\mathbf{Q}(t) \triangleq \begin{pmatrix} -\lambda(t;\mathbf{x}) & \lambda(t;\mathbf{x}) \\ \mu(t;\mathbf{x}) & -\mu(t;\mathbf{x}) \end{pmatrix},$$

where the rate functions

$$\begin{cases} \lambda(t;\mathbf{x}) \triangleq \frac{at^{a-1}}{1+t^a} \cdot e^{\mathbf{x}^\top \boldsymbol{\beta}_\lambda}, \\ \mu(t;\mathbf{x}) \triangleq \frac{bt^{b-1}}{1+t^b} \cdot e^{\mathbf{x}^\top \boldsymbol{\beta}_\mu}, \end{cases}$$

stand for the transition rates from state 0 to 1 and from state 1 to 0, respectively, at time $t \geq 0$, with a and b as the parameters for $\lambda(t;\mathbf{x})$ and $\mu(t;\mathbf{x})$, respectively. Denote $\mathbf{x} \in \mathbb{R}^d$ as the d covariates for each subject, and $\boldsymbol{\beta}_\lambda, \boldsymbol{\beta}_\mu \in \mathbb{R}^d$ as the corresponding coefficients for exponent factors of each state. These rate functions are adopted in the form of generalized log-logistic function due to its common application in survival analysis, and its flexibility for adapting both monotone and unimodal hazard functions.[38] Note the time-dependent nature of rate functions, while the rates for ordinary time-constant CTMC are denoted as λ and μ that do not depend on the time.

Let $Y(t) \in \{0,1\}$ be the state of the subject at time $t \geq 0$. We define the probability of transitioning states between times s and $s+t$ as $\mathbf{P}_{ij}(s,t;\mathbf{x}) \triangleq \mathbb{P}(Y(s+t) = j \mid Y(s) = i;\ \mathbf{x})$ for $0 \leq s \leq s+t$ and $i,j \in \{0,1\}$. Then, the generalized time-dependent probability transition matrix is

$$\mathbf{P}(s,t;\mathbf{x}) = \begin{pmatrix} \mathbf{P}_{00}(s,t;\mathbf{x}) & \mathbf{P}_{01}(s,t;\mathbf{x}) \\ \mathbf{P}_{10}(s,t;\mathbf{x}) & \mathbf{P}_{11}(s,t;\mathbf{x}) \end{pmatrix}.$$

The transition probabilities are derived from the following Chapman-Kolmogorov equations:

$$\mathbf{P}_{00}(s,t+\Delta t;\mathbf{x}) = \mathbf{P}_{00}(s,t;\mathbf{x})(1-\lambda(s+t;\mathbf{x})\Delta t + o(\Delta t)) + \mathbf{P}_{01}(s,t;\mathbf{x})(\mu(s+t;\mathbf{x})\Delta t + o(\Delta t)),$$

and

$$\mathbf{P}_{11}(s,t+\Delta t;\mathbf{x}) = \mathbf{P}_{11}(s,t;\mathbf{x})(1-\mu(s+t;\mathbf{x})\Delta t + o(\Delta t)) + \mathbf{P}_{10}(s,t;\mathbf{x})(\lambda(s+t;\mathbf{x})\Delta t + o(\Delta t)),$$

by obtaining the rates $\mathbf{P}'_{00}$ and $\mathbf{P}'_{11}$ of transition probabilities. Then, using the ordinary differential equations method, we obtain the following transition probabilities:

$$\begin{cases} \mathbf{P}_{00}(s,t;\mathbf{x}) = \frac{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}} u^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}-1}\,du + (1+s^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+s^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}{(1+(s+t)^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+(s+t)^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}, \\ \mathbf{P}_{01}(s,t;\mathbf{x}) = 1 - \frac{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}} u^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}-1}\,du + (1+s^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+s^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}{(1+(s+t)^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+(s+t)^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}, \\ \mathbf{P}_{10}(s,t;\mathbf{x}) = 1 - \frac{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}\int_{1+s^a}^{1+(s+t)^a}\left(1-(v-1)^{b/a}\right)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}} v^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}-1}\,dv + (1+s^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+s^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}{(1+(s+t)^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+(s+t)^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}, \\ \mathbf{P}_{11}(s,t;\mathbf{x}) = \frac{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}\int_{1+s^a}^{1+(s+t)^a}\left(1-(v-1)^{b/a}\right)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}} v^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}-1}\,dv + (1+s^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+s^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}{(1+(s+t)^a)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda}}(1+(s+t)^b)^{e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu}}}. \end{cases} \tag{1}$$

When we do not consider covariates so that $\mathbf{x} = \mathbf{0}$ and thereby $e^{\mathbf{x}^\top\boldsymbol{\beta}_\lambda} = e^{\mathbf{x}^\top\boldsymbol{\beta}_\mu} = 1$ for all subjects, the integrals are solvable and consequently give:

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\begin{cases} \mathbf{P}_{00}(s,t) = \dfrac{\frac{b\left((s+t)^{a+b}-s^{a+b}\right)}{a+b}+(s+t)^b-s^b+(1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}, \\ \mathbf{P}_{01}(s,t) = 1-\dfrac{\frac{b\left((s+t)^{a+b}-s^{a+b}\right)}{a+b}+(s+t)^b-s^b+(1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}, \\ \mathbf{P}_{10}(s,t) = 1-\dfrac{\frac{a\left((s+t)^{a+b}-s^{a+b}\right)}{a+b}+(s+t)^a-s^a+(1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}, \\ \mathbf{P}_{11}(s,t) = \dfrac{\frac{a\left((s+t)^{a+b}-s^{a+b}\right)}{a+b}+(s+t)^a-s^a+(1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}. \end{cases} \tag{2}$$

The derivations of these probabilities are explicitly shown in Appendix A.1.

3 | MAXIMUM LIKELIHOOD ESTIMATION

Once the model is established, we conducted statistical inferences by developing an algorithm to calculate the maximum likelihood estimates and validated this algorithm through simulations. In Sections 3 and 4, we formulated the likelihood function for the parameters of the interest, and then estimated the maximum likelihood estimators. We also developed a method for simulating the Markov chain and validating the model.

To obtain maximum likelihood estimators for the parameters a, b, $\boldsymbol{\beta}_\lambda$, and $\boldsymbol{\beta}_\mu$, the likelihood functions for the Markov chain are formulated. Assuming there are n subjects in a study, the likelihood for jth individual with m_j observations is the following:

$$\begin{aligned} L_j\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right) \triangleq \prod_{i=0}^{m_j-1} &\mathbf{P}_{00}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_0^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right)} \cdot \mathbf{P}_{01}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_0^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right)} \\ &\cdot \mathbf{P}_{10}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_1^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right)} \cdot \mathbf{P}_{11}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_1^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right)}, \end{aligned}$$

where $\mathbf{x}_j \in \mathbb{R}^d$ stands for the covariates of jth individual. The likelihood is computed through observing the Markov chain from for time $t = t_0^{(j)}, \ldots, t_{m_j}^{(j)}$, with the indicator function for individual j being in state k at time t:

$$\mathbb{I}_k^{(j)}(t) \triangleq \begin{cases} 1 & \text{if } Y_j(t) = k, \\ 0 & \text{otherwise}, \end{cases}$$

for $k \in \{0,1\}$, where $Y_j\left(t_i^{(j)}\right) \in \{0,1\}$ represents the state of jth individual at observation time point $t_i^{(j)}$. When we consider all n subjects so that $j = 1, \ldots, n$, the full likelihood is updated to the following form:

$$\begin{aligned} L\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right) \triangleq \prod_{j=1}^{n} L_j\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right) = \prod_{j=1}^{n}\prod_{i=0}^{m_j-1} &\mathbf{P}_{00}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_0^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right)} \\ &\cdot \mathbf{P}_{01}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_0^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right)} \cdot \mathbf{P}_{10}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_1^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right)} \\ &\cdot \mathbf{P}_{11}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right)^{\mathbb{I}_1^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right)}. \end{aligned}$$

Therefore, the log-likelihood, the function of interest for maximization, is the following:

$$\begin{aligned} l\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right) &\triangleq \log\left(L\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right)\right) = \log\left(\prod_{j=1}^{n} L_j\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right)\right) \equiv \sum_{j=1}^{n} l_j\left(a,b,\boldsymbol{\beta}_\lambda,\boldsymbol{\beta}_\mu\right) \\ &= \sum_{j=1}^{n}\sum_{i=0}^{m_j-1} \mathbb{I}_0^{(j)}\left(t_i^{(j)}\right)\cdot\mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right)\cdot\log\mathbf{P}_{00}\left(t_i^{(j)}, t_{i+1}^{(j)}-t_i^{(j)};\mathbf{x}_j\right) \end{aligned}$$

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$+ \mathbb{I}_0^{(j)}\left(t_i^{(j)}\right) \cdot \mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right) \cdot \log \mathbf{P}_{01}\left(t_i^{(j)}, t_{i+1}^{(j)} - t_i^{(j)}; \mathbf{x}_j\right)$$
$$+ \mathbb{I}_1^{(j)}\left(t_i^{(j)}\right) \cdot \mathbb{I}_0^{(j)}\left(t_{i+1}^{(j)}\right) \cdot \log \mathbf{P}_{10}\left(t_i^{(j)}, t_{i+1}^{(j)} - t_i^{(j)}; \mathbf{x}_j\right)$$
$$+ \mathbb{I}_1^{(j)}\left(t_i^{(j)}\right) \cdot \mathbb{I}_1^{(j)}\left(t_{i+1}^{(j)}\right) \cdot \log \mathbf{P}_{11}\left(t_i^{(j)}, t_{i+1}^{(j)} - t_i^{(j)}; \mathbf{x}_j\right). \quad (3)$$

For the non-covariate cases, $\mathbf{x}_j = 0$ for all j, so that $\mathbf{P}_{ij}(s, t; \mathbf{x})$ in the log-likelihood (3) are replaced with $\mathbf{P}_{ij}(s, t)$ in (2).

4 | SIMULATIONS ANALYSIS

The process of the continuous-time Markov chain prior to switching to a new state resembles that of the Poisson process. As a result, based on the simulation of the non-homogeneous Poisson process, a similar algorithm was used to simulate the process of a two-state NH-CTMC. In this section, we illustrated the algorithms for simulating each state, simulating the whole Markov chain, and estimating maximum likelihood estimates for parameters using the pre-defined likelihood functions in Section 3, and the corresponding results.

4.1 | Non-homogeneous Poisson process simulation by thinning process

The transition from state 0 to state 1 can be considered as the first event of the non-homogeneous Poisson process with a rate of $\lambda(t; \mathbf{x})$, as well as the transition from state 1 to 0, with a rate of $\mu(t; \mathbf{x})$. Using this idea, we can simulate the event time of switching states by taking the first event time of the respective non-homogeneous Poisson process.

Non-homogeneous Poisson process can be simulated by using thinning process, which is explicitly elaborated and validated by Lewis et al[39] and Saltzman et al.[40] For simulating a *homogeneous* Poisson process, one of the possible ways is to simulate the number of events with Poisson distribution and then distribute the simulated events through time interval; the arrival times of events follow *i.i.d.* uniform distribution. However, for the non-homogeneous Poisson process, the arrival times are not distributed uniformly due to the time-variant nature of the rate. Hence, thinning process intervenes to replace uniform distribution and distribute the events accordingly to the time-specific rates.

State transitions of the conventional time-independent CTMC can be viewed as Poisson process events; the underlying assumptions of Poisson process are independence of events occurrences and stationarity of rate. Poisson distribution of the number of events and exponential distribution of interarrival times are the byproducts of independence and stationarity. While NH-CTMC is also based on the independence assumption, the stationarity does not hold any more due to time-dependence of the rate function. Hence the state transitions can be rather considered as non-homogeneous Poisson process events, and the corresponding simulation involves thinning process.

The simulation was based on the Algorithm 1. The events were simulated as homogeneous Poisson process with rate at the maximum value of the rate function $\lambda(t)$. The thinning happened afterwards, where the events were determined to be kept or dropped based on the probability computed by each event's rate divided by the maximum rate. With this simulation tool, we elaborated how to simulate the entire Markov chain in the next section.

4.2 | Simulation of two-state NH-CTMC

Using the thinning process of the non-homogeneous Poisson process simulation, we developed the algorithm to simulate the two-state NH-CTMC. The simulation algorithm is detailed in Algorithm 2. The idea is to pick the first event of non-homogeneous Poisson process and consider it as the state transition event. The simulation continued until it reached the endpoint time T; note that there might be no transition throughout the remaining time, indicating that the individual would stay in the single, initial state until the end. In this simulation, $n = 100$ subjects were simulated from time 0 to $T = 50$ with single-dimension covariates, and simulations were conducted for 1000 replications to compare and analyze the results. Note that the choice of sample size and observational duration is comparable with the DASH data illustrated in Section 5. Figure 1 shows the rate functions $\lambda(t; x)$ and $\mu(t; x)$, both in the form of the log-logistic baseline hazard function with and without covariates. The covariate x_j and the initial state $Y_j(0)$ of each jth subject followed the binary distribution with identical probabilities. When the simulation was completed for all the individuals up to time T, the data

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

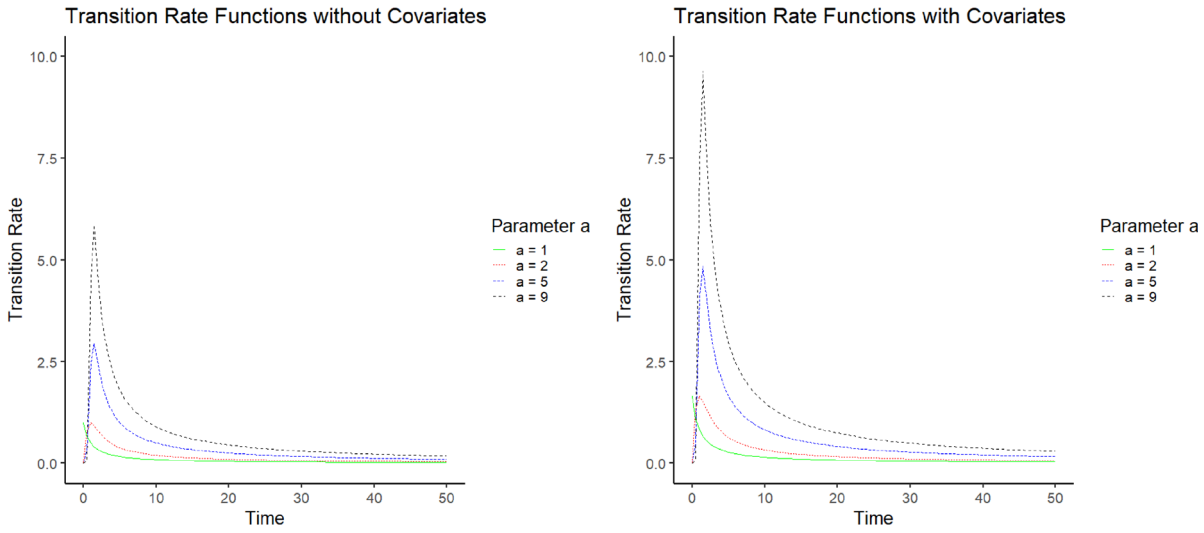

FIGURE 1 Plot of rate functions with different rates, without (left) and with (right) covariates, for $t \in [0, 50]$ and $\beta = 0.2$

were collected on the discrete time points $t = 0, 1, \ldots, T-1, T$. The likelihood was estimated based on the Equation (3) on the simulated dataset, and the maximum likelihood estimates were computed for each run.

With the simulated dataset, the log-likelihood function in Section 3 for maximum likelihood estimation could be formed for the parameters a, b, β_λ, and β_μ. For the estimation, we used the function `mle2` from the `R` package `bbmle`.[41] The optimizer used in the function `mle2` was `Nelder-Mead`,[42] a gradient-free optimization method where constraints are formulated to ensure the feasibility of $a, b > 0$ throughout the heuristic process. Note that β_λ and β_μ are not binded to constraints. Since the Nelder-Mead method is a gradient-free optimization method with heuristic simplex method, it is more suitable for complex functions which may not have a closed form gradient. Consequently, the method works relatively efficiently on lower-dimension functions.[43] In this case, the log-likelihood function $l : \mathbb{R}^{2(d+1)} \mapsto \mathbb{R}$ had four parameters present, as there was only one covariate, that is, $d = 1$. We set $a_0 = b_0 = 1$ and $\beta_\lambda^{(0)} = \beta_\mu^{(0)} = 0$ for initial estimates of parameters for `mle2` procedure.

4.3 | Results

The simulation examined a total of six scenarios. First, in the case without covariates, two sets of parameters for the transition rates were considered: $(a, b)^\top = (9, 5)^\top$ and $(2, 1)^\top$. These correspond to situations with fast and slow transition rates respectively. In the case where a covariate was present, we considered two sets of coefficients with a single dimension to see if the model was able to estimate both larger and smaller covariate effects: $(\beta_\lambda, \beta_\mu)^\top = (1.0, 0.5)^\top$ and $(0.2, -0.1)^\top$. The results were summarized in Table 1, where percentage of bias (PB) is computed by dividing bias over the true parameter value, and coverage probability (CP) stands for the proportion of 95% confidence intervals ($Z_{\alpha/2} = 1.96$) that cover the true parameter among the entire simulation runs. Figure 2 shows the receiver operating characteristic (ROC) curves of simulations with six different parameter sets. The area under curve (AUC) and accuracy metrics were obtained from the observed states and their predicted values of all the samples, through 10-fold cross-validation process on 1000 simulated samples. In Figure 3, longitudinal logistic regression models were fitted onto the simulation data in order to compare the performances to those of NH-CTMC models. The overall results demonstrate the stability and superiority of the NH-CTMC model, showing over 93% CP's (coverage probabilities) for the parameters and low biases as in Table 1, and high AUC's and accuracies as in Table 2.

5 | APPLICATION TO HYPERTENSION DATA

The DASH study is a multicenter randomized controlled trial, with the goal of determining the effects of dietary patterns in reducing blood pressure.[7] The DASH dataset measured 24-h ambulatory blood pressure for 348 patients, with observations

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 The simulation results with 1000 simulation runs, with $n = 100$ samples and up to time unit $T = 50$, listing the bias, percentage of bias (PB), standard deviation of estimates (SD), square root of the mean of the squared error (MSE), and percent coverage probability (CP)

Without covariates		**True parameter**	**Bias**	**PB (%)**	**SD**	**SE**	**MSE**	**CP (%)**
1.	a	9	0.0140	0.1558%	0.4452	0.4432	0.1982	95.10%
	b	5	−0.0074	0.1483%	0.2380	0.2448	0.0566	96.50%
2.	a	2	0.0204	1.0195%	0.1376	0.1389	0.0193	95.10%
	b	1	−0.0033	0.3297%	0.0673	0.0651	0.0045	92.90%
With covariates		**True parameter**	**Bias**	**PB (%)**	**SD**	**SE**	**MSE**	**CP (%)**
3.	a	9	0.0262	0.2907%	0.6430	0.6276	0.4137	95.20%
	b	5	0.0103	0.2052%	0.3579	0.3499	0.1281	93.30%
	β_λ	1.0	−0.0002	0.0241%	0.1025	0.1007	0.0105	94.70%
	β_μ	0.5	0.0003	0.0579%	0.1022	0.1032	0.0104	95.10%
4.	a	9	0.0848	0.9426%	0.6389	0.6254	0.4150	94.90%
	b	5	0.0311	0.6211%	0.3580	0.3480	0.1290	94.20%
	β_λ	0.2	−0.0076	3.8218%	0.1000	0.0969	0.0101	94.00%
	β_μ	−0.1	−0.0032	3.2018%	0.1019	0.0993	0.0104	94.20%
5.	a	2	0.0278	1.3898%	0.2027	0.1938	0.0418	94.10%
	b	1	0.0040	0.4022%	0.0917	0.0928	0.0084	95.10%
	β_λ	1.0	−0.0037	0.3743%	0.1332	0.1303	0.0177	93.90%
	β_μ	0.5	0.0017	0.3363%	0.1422	0.1408	0.0202	95.00%
6.	a	2	0.0156	0.7800%	0.1877	0.1901	0.0354	94.90%
	b	1	0.0084	0.8397%	0.0915	0.0923	0.0084	95.20%
	β_λ	0.2	0.0006	0.3065%	0.1327	0.1382	0.0176	95.80%
	β_μ	−0.1	−0.0055	5.4748%	0.1546	0.1502	0.0239	94.70%

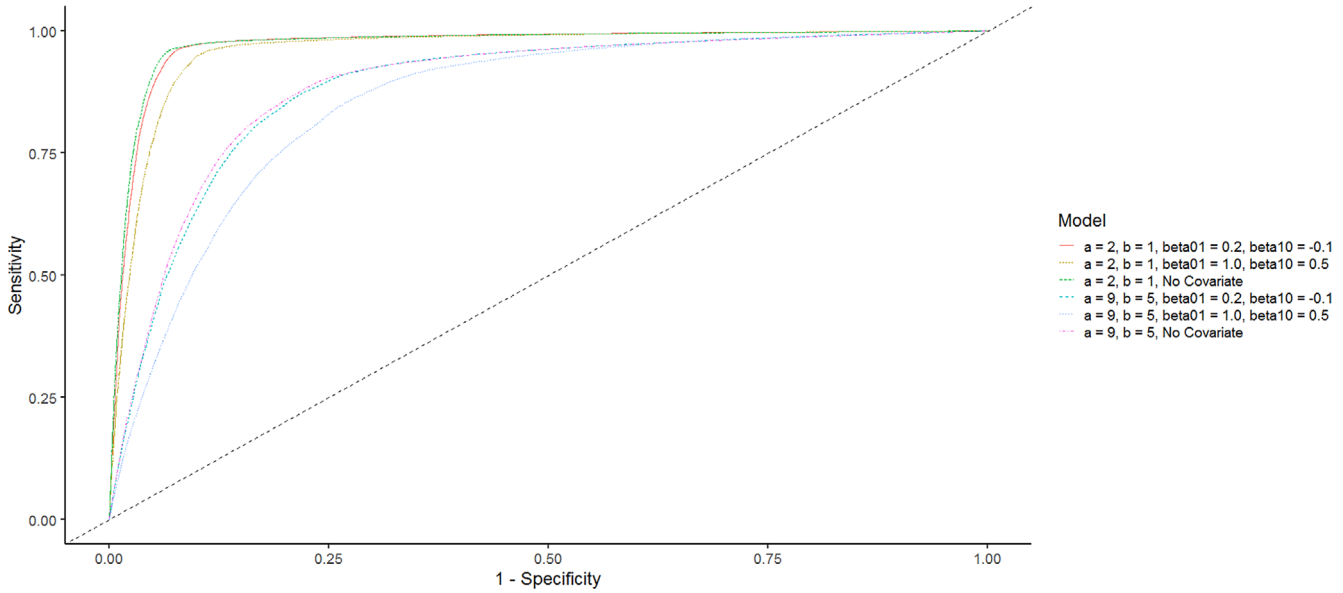

FIGURE 2 ROC curves of NH-CTMC models on simulated datasets with six different parameter sets, with AUC and accuracy results as in Table 2

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

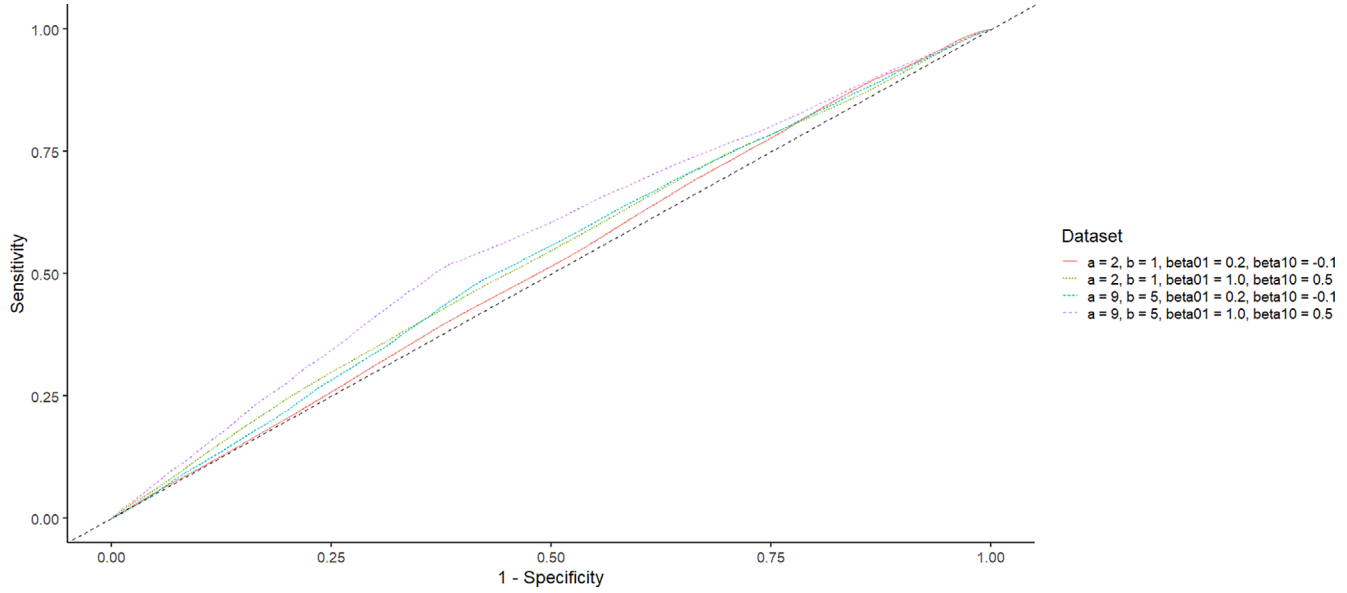

FIGURE 3 ROC curves of longitudinal logistic regression models on simulated datasets with four different parameter sets, with AUC and accuracy results as in Table 2

TABLE 2 The receiver operating characteristic (ROC) curves results of 10-fold cross validation of NH-CTMC model and longitudinal logistic regression models on different simulated datasets, listing the area under curve (AUC) and accuracy

Parameter				NH-CTMC		Longitudinal logit	
a	b	β_λ	β_μ	AUC	Accuracy	AUC	Accuracy
2	1	0.2	−0.1	0.9710	95.22%	0.5167	71.46%
2	1	1.0	0.5	0.9604	93.91%	0.5350	71.04%
2	1	0	0	0.9732	95.12%	-	-
9	5	0.2	−0.1	0.8855	85.23%	0.5339	67.00%
9	5	1.0	0.5	0.8554	83.16%	0.5701	68.70%
9	5	0	0	0.8889	84.97%	-	-

recorded every 30 min on average. Each record included values of systolic blood pressure (SBP), time of observation, and covariates information including dietary plan. The data was extracted for each subject with the start time $t = 0$ defined as the time when the subject woke up.

The state of hypertension, measurement time, and additional covariates were derived from the dataset. The state of hypertension was based on SBP measures; a threshold of 130 mmHg in SBP was applied to define the binary states of hypertension.[44] Measurement time was calculated as the elapsed time from the patient waking up, with units in hour. Four covariates were taken into account for the modeling that were studied to have effects on hypertension: dietary plan, gender, age group, and BMI. Patients were initially divided into three dietary interventions: the control diet, an intermediate diet, and a reduced diet with lower fat contents and higher vegetables, fruits, proteins, and dairy products. The patients with intermediate diet control had dietary plan similar to the control diet group overall, except for having higher amounts of fruits and vegetables. We considered dietary intervention as a binary covariate where the first group (0) consists of the control and intermediate diets, and the second group (1) was assigned the reduced diet. Gender was assigned as 1 for male and 0 for female; age group was dichotomized into ≤ 50 years old (0) and > 50 years old or above (1); BMI, calculated from weight and height, was taken into account as a continuous covariate.

The overall performance of our model was evaluated using 10-fold cross validation. For each iteration, $\hat{a}$, $\hat{b}$, $\hat{\boldsymbol{\beta}}_\lambda$, and $\hat{\boldsymbol{\beta}}_\mu$ were estimated from training data then applied to compute the hypertension probability of test data, that is,

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Parameter estimation results for NH-CTMC model on DASH study, with the covariates of dietary plan, sex, age group, and BMI, respectively, at the non-hypertensive states ($\boldsymbol{\beta}_\lambda$) and hypertensive ($\boldsymbol{\beta}_\mu$) states

Parameter	Estimate	Standard error	Lower 95% CI	Upper 95% CI
a	2.0473	0.1499	1.7535	2.3412
b	1.9407	0.1566	1.6337	2.2477
$\beta_\lambda^{\text{diet}}$	−0.3012	0.0823	−0.4626	−0.1398
$\beta_\lambda^{\text{sex}}$	0.1607	0.0370	0.0882	0.2333
$\beta_\lambda^{\text{age}}$	0.2266	0.0675	0.0943	0.3590
$\beta_\lambda^{\text{BMI}}$	0.0044	0.0011	0.0023	0.0065
β_μ^{diet}	−0.1031	0.0824	−0.2646	0.0584
β_μ^{sex}	0.0781	0.0766	−0.0721	0.2283
β_μ^{age}	−0.6558	0.0915	−0.8352	−0.4764
β_μ^{BMI}	0.0139	0.0036	0.0069	0.0210

TABLE 4 Parameter estimation results for longitudinal logistic regression model on DASH study, with the covariates of recording time elapsed, dietary plan, sex, age group, and BMI

Parameter	Estimate	Standard error	Lower 95% CI	Upper 95% CI
θ_{time}	0.1261	0.0180	0.0908	0.1614
θ_{diet}	−0.3529	0.2544	−0.8514	0.1457
θ_{sex}	0.3321	0.2416	−0.1414	0.8056
θ_{age}	1.3899	0.2664	0.8679	1.9119
θ_{BMI}	−0.0574	0.0327	−0.1216	0.0067

$\mathbf{P}_{i1}(\cdot)$, $i \in \{0, 1\}$, of each time-point observation of a subject. Based on the prediction result, the areas under receiver operating characteristic curves (AUROC) and the accuracy rates were computed. The estimates of NH-CTMC parameters and longitudinal logistic regression parameters were summarized in Tables 3 and 4, respectively, with their standard errors and 95% confidence intervals. Figure 4 shows that the NH-CTMC model outperformed the longitudinal logistic regression model with respect to both AUC and accuracy metrics, improving both from near 60% to above 75%.

The NH-CTMC model also produced results with more intuitive interpretability; the MLE estimate $\hat{a} = 2.0473$ of parameter for the rate function λ of control dietary group had a negligibly higher value than the estimate $\hat{b} = 1.9407$ of the rate function μ of case dietary group, indicating that the transition towards hypertension state happened slightly faster than the transition towards low SBP state for the in-patients. In addition, the $\hat{\boldsymbol{\beta}}$'s provided information on the effect of the covariates on hypertension; the most statistically significant covariate was age ($\beta_\lambda^{(\text{age})}, \beta_\mu^{(\text{age})}$). The positive beta values $\beta_\lambda^{(\text{age})} = 0.2266$ and $\theta_{\text{age}} = 1.3899$ from NH-CTMC and logit models, respectively, implied that the transitions towards high SBP state were faster for the group with ages of greater than 50 years old. The negative beta value $\beta_\mu^{(\text{age})} = -0.6558$ from NH-CTMC model similarly indicated that the elderly group had a slower rate on the transitions towards non-hypertensive state. While diet group and sex were insignificant covariates from the longitudinal logistic model, they were significant covariates for transitioning from non-hypertensive state to hypertensive state in the NH-CTMC model. $\beta_\lambda^{(\text{diet})} = -0.3012$ indicated that the assigned reduced diet was able to decelerate the transition into hypertensive state, and $\beta_\lambda^{(\text{sex})} = 0.1607$ implied an accelerated transition into hypertensive state among males. On the other hand, BMI was not significant in the logistic model, and the small estimates from the NH-CTMC model infers that its effect could be negligible.

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

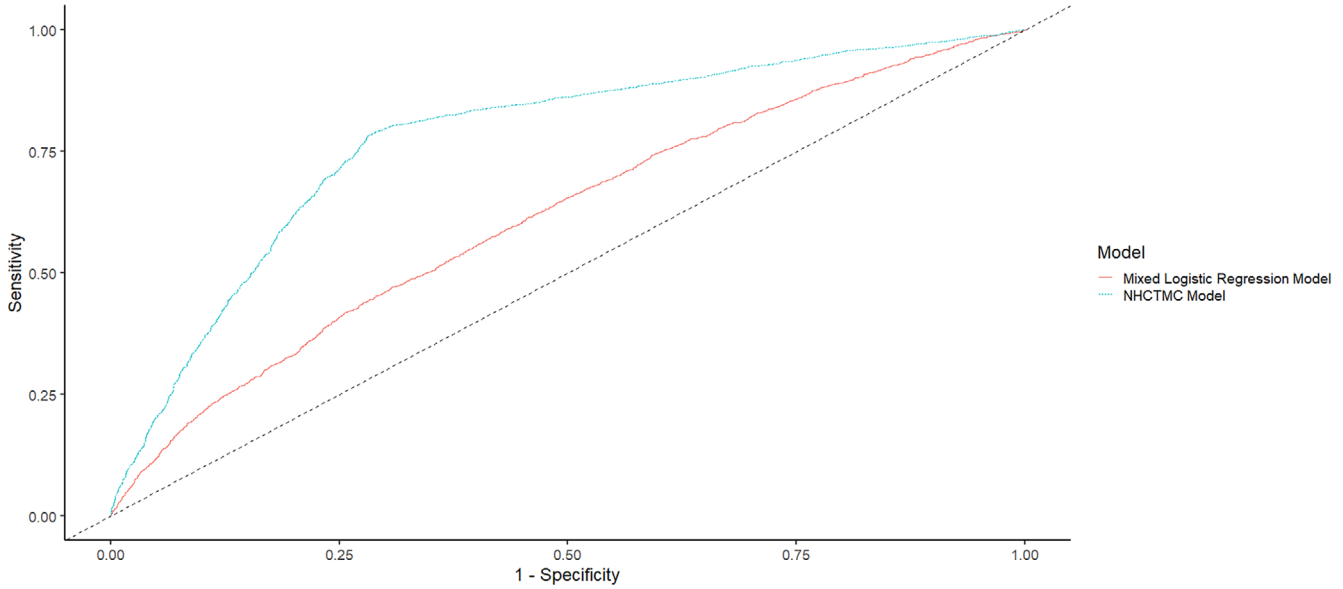

FIGURE 4 ROC curves of longitudinal logistic regression and NH-CTMC models on DASH data, with AUC of 0.6149 and 0.7737, respectively, and accuracy of 60.28% and 75.69%, respectively

6 | DISCUSSION

Continuous-time Markov chains over finite state-spaces are widely used to model dynamic processes in many natural and social science fields. In this study, the closed-form solution for a two-state NH-CTMC was derived and confirmed through simulations and application to a real dataset. The details of the derivation were presented in Appendix A. The computational results show the strength and effectiveness of the process. The application to DASH study demonstrated the potential for the suggested methodology to be used in future medical research. To the best of our knowledge, this is the first attempt for having a closed form of transition probabilities for two-state CTMC with time-variant rates and covariates.

The performance of NH-CTMC model was verified by the simulation results at Section 4. Our approach has an advantage in capturing the time-variant rate of the data, and the prediction accuracy remained high for both high and low pairs of parameters, namely compared to the longitudinal logistic regression model. One minor point to notice is that both the AUC and accuracy for the parameter set with lower values were relatively higher than those for the parameter set with higher values. Since the transitions occurred more frequently for higher-valued parameters within one unit time, some of the transitions were lost when the data were sampled through discrete time points. Nevertheless, the simulation results showed the well-fitting behavior of NH-CTMC model on longitudinal data with the proposed non-homogeneous rate.

The NH-CTMC model has several advantages compared to other conventional longitudinal models. In comparison to the traditional CTMC, NH-CTMC is able to capture the varying rate of state transition event and oversee the trend of it as in Figure 1. The parameters of transition rates represent the frequency of transitions happening, while $\boldsymbol{\beta}$'s amplify or reduce the transition rate, depending on their signs and presence of the covariates. Based on the estimate of $\boldsymbol{\beta}$'s, we are able to make inference on the effect of covariates on the process. Such modeling and inference on covariates are not feasible in conventional CTMC, thereby differentiating the NH-CTMC in a complex longitudinal setting. Compared to the longitudinal logistic regression model, NH-CTMC is able to retain the process' transitioning nature due to its Markovian property. Although the longitudinal logistic regression model is less computation intensive to estimate, Table 2 and Figures 2–4 demonstrate the reduction in accuracy when ignoring the transition between the states of outcome. Thus, NH-CTMC is able to provide insight into the transitions between different time points in the longitudinal framework while capturing time-variant rates in the process. Due to its flexibility, NH-CTMC is advantageous for many settings in medical research.

There are some limitations for the NH-CTMC model. First, the model is defined by a specific form of transition rate which may not be applicable to some longitudinal processes. Second, the maximum likelihood estimation is computationally intensive for the covariates case due to the integrals in the transition probabilities which must be

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Algorithm 1. Non-homogeneous Poisson process simulation using thinning process

Require: Rate function $\lambda(t)$, simulation time $T > 0$

$\lambda^*(T) \triangleq \max_{\tau \in [0,T]} \lambda(\tau)$
Simulate the number of events: $n \sim Poisson\,(\lambda^*(T))$
Simulate the events' times: $t_j \sim U(0, T)$ for $j = 1, \dots, n$
for $j = 1, \dots, n$ **do**
    $p_j \triangleq \frac{\lambda(t_j)}{\lambda^*(T)}$
    Simulate $X_j \sim Bernoulli(p_j)$
    if $X_j = 1$ **then**
        Keep the event at t_j for the simulation.
    else
        Drop the event at t_j.
    end if
end for

Algorithm 2. Two-state NH-CTMC simulation for individual subject

Require: Rate functions $\lambda(t; x)$ & $\mu(t; x)$, parameters of rate functions $a, b, \beta_\lambda, \beta_\mu$, simulation time range $T > 0$, covariate x, initial state $Y(0) \in \{0, 1\}$, stopping parameter $\varepsilon > 0$

$t \leftarrow 0$
while $t \leq T$ **do**
    $f(t, \tau; x) \triangleq \lambda(\tau; x) \cdot \mathbb{I}_0(t) + \mu(\tau; x) \cdot \mathbb{I}_1(t)$ ▷ $\mathbb{I}_k(t) \triangleq \begin{cases} 1 & \text{if } Y(t) = k, \\ 0 & \text{otherwise.} \end{cases}$
    $F^*(t, T; x) \triangleq \max_{\tau \in [t,T]} f(t, \tau; x)$ ▷ Maximum of the rate at time interval $[t, T]$
    Simulate $n_t \sim Poisson\,(F^*(t, T; x))$ ▷ The number of transition events.
    if $n_t = 0$ **then** ▷ No transition is happening until time T.
        $t \leftarrow T + \varepsilon$, and go to line 22. ▷ Terminate the simulation.
    end if
    Simulate $t_j \sim U(t, T), j = 1, \dots, n_t$. WLOG, assume $t_1 \leq t_2 \leq \cdots \leq t_{n_t}$
    $k \leftarrow 0, X_k \leftarrow 0$
    while $k < n_t$ and $X_k = 0$ **do**
        $k \leftarrow k + 1$
        Simulate $X_k \sim Bernoulli(p_{k,x})$, where $p_{k,x} \triangleq \frac{f(t,t_k;x)}{F^*(t,T;x)}$
        if $X_k = 1$ **then** ▷ Inner while loop will terminate.
            $Y(t_k) \leftarrow 1 - Y(t)$ ▷ Transition takes place.
            $t \leftarrow t_k$
        end if
    end while
    if $X_k = 0$ **then** ▷ A case where all of the transition events were dropped.
        $t \leftarrow T + \varepsilon$ ▷ Terminate the simulation.
    end if
end while

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

numerically evaluated. Hence, the future directions of this work are as follows: (a) consider transition probabilities with other forms of rate functions, such as Bernstein polynomial and normalized/resetting time units, which could be derived through similar procedures, (b) improve the model by proposing a general form of rate functions, while keeping the corresponding probability functions as simple as possible, and (c) further simplify the current NH-CTMC model by obtaining a closed approximation of the covariates integral.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the reviewers for their helpful and insightful comments, which served to strengthen this work.

DATA AVAILABILITY STATEMENT

Data used in the preparation of this article was obtained from the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC). The objective of the Dietary Approaches to Stop Hypertension (DASH) study was to test the effects of dietary patterns on reducing hypertension. To request data gathered from the DASH study, visit https://biolincc.nhlbi.nih.gov/studies/dash/.

ORCID

Joonha Chang https://orcid.org/0000-0002-2343-8465
Hei Kit Chan https://orcid.org/0000-0001-7648-4126
Jeffrey Lin https://orcid.org/0000-0003-0180-0117
Wenyaw Chan https://orcid.org/0000-0003-3329-5282

REFERENCES

1. Sidney S, Rosamond WD, Howard VJ, Luepker RV. The "heart disease and stroke statistics—2013 update" and the need for a national cardiovascular surveillance system. *Am Heart Assoc*. 2013;127:21-23.
2. George J, MacDonald T. Home blood pressure monitoring. *Eur Cardiol Rev*. 2015;10(2):95.
3. Jain V, Minhas AMK, Morris AA, et al. Demographic and regional trends of heart failure–related mortality in young adults in the US, 1999-2019. *JAMA Cardiol*. 2022;7(9):900-904.
4. Stierman B, Afful J, Carroll MD, et al. National Health and Nutrition Examination Survey 2017–March 2020 prepandemic data files development of files and prevalence estimates for selected health outcomes. National Health Statistics Reports; 2021:158.
5. Ritchey MD, Loustalot F, Bowman BA, Hong Y. Trends in mortality rates by subtypes of heart disease in the United States, 2000-2010. *JAMA*. 2014;312(19):2037-2039.
6. Lu Y, Lan T. Global, regional, and national burden of hypertensive heart disease during 1990–2019: an analysis of the global burden of disease study 2019. *BMC Public Health*. 2022;22(1):1-10.
7. Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. *N Engl J Med*. 1997;336(16):1117-1124.
8. Gillis EE, Sullivan JC. Sex differences in hypertension: recent advances. *Hypertension*. 2016;68(6):1322-1327.
9. Buford TW. Hypertension and aging. *Ageing Res Rev*. 2016;26:96-111.
10. Sabaka P, Dukat A, Gajdosik J, Bendzala M, Caprnda M, Simko F. The effects of body weight loss and gain on arterial hypertension control: an observational prospective study. *Eur J Med Res*. 2017;22(1):1-7.
11. Vasudevan A, Thomas T, Kurpad A, Sachdev HS. Prevalence of and factors associated with high blood pressure among adolescents in India. *JAMA Netw Open*. 2022;5(10):e2239282.
12. Islam SMS, Talukder A, Awal MA, et al. Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries. *Front Cardiovasc Med*. 2022;9:839379.
13. Rossi GP, Cesari M, Cuspidi C, et al. Long-term control of arterial hypertension and regression of left ventricular hypertrophy with treatment of primary aldosteronism. *Hypertension*. 2013;62(1):62-69.
14. LaFreniere D, Zulkernine F, Barber D, Martin K. Using machine learning to predict hypertension from a clinical dataset. Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE; 2016:1–7.
15. Suzuki Y, Kaneko H, Yano Y, et al. Association of cardiovascular health metrics with risk of transition to hypertension in non-hypertensive young adults. *Am J Hypertens*. 2022;35(10):858-866. doi:10.1093/ajh/hpac057
16. Tsimihodimos V, Gonzalez-Villalpando C, Meigs JB, Ferrannini E. Hypertension and diabetes mellitus. *Hypertension*. 2018;71(3):422-428. https://www.ahajournals.org/doi/abs/10.1161/HYPERTENSIONAHA.117.10546
17. Chen HH, Duffy SW, Tabar L. A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. *J R Stat Soc Ser D Stat*. 1996;45(3):307-317.
18. Combescure C, Chanez P, Saint-Pierre P, et al. Assessment of variations in control of asthma over time. *Eur Respir J*. 2003;22(2):298-304.
19. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Stat Med*. 1995;14(14):1531-1543.

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

20. Gentleman RC, Lawless JF, Lindsey JC, Yan P. Multi-state Markov models for analysing incomplete disease history data with illustrations for HIV disease. *Stat Med*. 1994;13(8):805-821.
21. Uhry Z, Hédelin G, Colonna M, et al. Multi-state Markov models in cancer screening evaluation: a brief review and case study. *Stat Methods Med Res*. 2010;19(5):463-486.
22. Kristal AR, Glanz K, Tilley BC, Li S. Mediating factors in dietary change: understanding the impact of a worksite nutrition intervention. *Health Educ Behav*. 2000;27(1):112-125.
23. Chen THH, Yen MF, Lai MS, et al. Evaluation of a selective screening for colorectal carcinoma: the Taiwan multicenter cancer screening (TAMCAS) project. *Cancer*. 1999;86(7):1116-1128.
24. Li YP, Chan W. Analysis of longitudinal multinomial outcome data. *Biom J*. 2006;48(2):319-326.
25. Ma J, Chan W, Tsai CL, Xiong M, Tilley BC. Analysis of transtheoretical model of health behavioral changes in a nutrition intervention study—a continuous time Markov chain model with Bayesian approach. *Stat Med*. 2015;34(27):3577-3589.
26. Mhoon KB, Chan W, Del Junco DJ, Vernon SW. A continuous-time Markov chain approach analyzing the stages of change construct from a health promotion intervention. *JP J Biostat*. 2010;4(3):213.
27. Ma J, Chan W, Tilley BC. Continuous time Markov chain approaches for analyzing transtheoretical models of health behavioral change: a case study and comparison of model estimations. *Stat Methods Med Res*. 2018;27(2):593-607.
28. Yang J, Liu F, Wang B, et al. Blood pressure states transition inference based on multi-state Markov model. *IEEE J Biomed Health Inform*. 2021;25(1):237-246.
29. Zheng X, Xiong J, Zhang Y, et al. Multistate Markov model application for blood pressure transition among the Chinese elderly population: a quantitative longitudinal study. *BMJ Open*. 2022;12(7):e059805. https://bmjopen.bmj.com/content/12/7/e059805
30. Wu Y, Wu W, Lin Y, Xiong J, Zheng X. Blood pressure states transitions among bus drivers: the application of multi-state Markov model. *Int Arch Occup Environ Health*. 2022;95:1995-2003.
31. Manios E, Tsagalis G, Tsivgoulis G, et al. Time rate of blood pressure variation is associated with impaired renal function in hypertensive patients. *J Hypertens*. 2009;27(11):2244-2248.
32. Smolensky MH, Hermida RC, Portaluppi F. Circadian mechanisms of 24-hour blood pressure regulation and patterning. *Sleep Med Rev*. 2017;33:4-16.
33. Johnson JT, Luecke GR. Nonhomogeneous, continuous-time Markov chains defined by series of proportional intensity matrices. *Stoch Process Their Appl*. 1989;32(1):171-181.
34. Zeifman A, Satin Y, Kovalev I, Razumchik R, Korolev V. Facilitating numerical solutions of inhomogeneous continuous time Markov chains using ergodicity bounds obtained with logarithmic norm method. *Mathematics*. 2020;9(1):42.
35. Reznícek J, Kohlík M, Kubátová H. Accurate inexact calculations of non-homogeneous Markov chains. Proceedings of the 2019 22nd Euromicro Conference on Digital System Design (DSD). IEEE; 2019:470–477.
36. Řezníček J, Kohlík M, Kubátová H. Non-homogeneous continuous time Markov chains calculations. Proceedings of the 2020 23rd Euromicro Conference on Digital System Design (DSD). IEEE; 2020:664–671.
37. Chen B, Zhou XH. Non-homogeneous Markov process models with informative observations with an application to Alzheimer's disease. *Biom J*. 2011;53(3):444-463.
38. Khan SA, Khosa SK. Generalized log-logistic proportional hazard model with applications in survival analysis. *J Stat Distrib Appl*. 2016;3:1-18.
39. Lewis PAW, Shedler GS. Simulation of nonhomogeneous Poisson processes by thinning. *Nav Res Logist Q*. 1979;26(3):403-413.
40. Saltzman EA, Drew JH, Leemis LM, Henderson SG. Simulating multivariate nonhomogeneous Poisson processes using projections. *ACM Trans Model Comput Simul*. 2012;22(3):1-13.
41. Bolker B, R Development Core Team. bbmle: tools for general maximum likelihood estimation. R package version 1.0.24; 2021. https://CRAN.R-project.org/package=bbmle
42. Nelder JA, Mead R. A simplex method for function minimization. *Comput J*. 1965;7(4):308-313.
43. Han L, Neumann M. Effect of dimensionality on the Nelder-Mead simplex method. *Optim Methods Softw*. 2006;21(1):1-16.
44. Carey RM, Whelton PK. Prevention, detection, evaluation, and management of high blood pressure in adults: synopsis of the 2017 American College of Cardiology/American Heart Association hypertension guideline. *Ann Internal Med*. 2018;168(5):351-358.

How to cite this article: Chang J, Chan HK, Lin J, Chan W. Non-homogeneous continuous-time Markov chain with covariates: Applications to ambulatory hypertension monitoring. *Statistics in Medicine*. 2023;1-16. doi: 10.1002/sim.9707

APPENDIX A

A.1 Derivations of transition probabilities

Since the transition probability $\mathbf{P}_{01}(s,t;\mathbf{x}) = 1 - \mathbf{P}_{00}(s,t;\mathbf{x})$ is automatically derived once $\mathbf{P}_{00}(s,t;\mathbf{x})$ is determined, we elaborate the derivation process of $\mathbf{P}_{00}(s,t;\mathbf{x})$. Consider the following Chapman-Kolmogorov equation:

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\mathbf{P}_{00}(s,t+\Delta t;\mathbf{x}) = \mathbf{P}_{00}(s,t;\mathbf{x})(1-\lambda(s+t;\mathbf{x})\Delta t + o(\Delta t)) + \mathbf{P}_{01}(s,t;\mathbf{x})(\mu(s+t;\mathbf{x})\Delta t + o(\Delta t)),$$

which can be rewritten as:

$$\frac{\mathbf{P}_{00}(s,t+\Delta t;\mathbf{x}) - \mathbf{P}_{00}(s,t;\mathbf{x})}{\Delta t} = \mathbf{P}_{01}(s,t;\mathbf{x})\mu(s+t;\mathbf{x}) - \mathbf{P}_{00}(s,t;\mathbf{x})\lambda(s+t;\mathbf{x}) + \frac{o(\Delta t)}{\Delta t}.$$

The differential equation can be derived by taking limit $\Delta t \to 0$:

$$\mathbf{P}'_{00}(s,t;\mathbf{x}) + \lambda(s+t;\mathbf{x})\mathbf{P}_{00}(s,t;\mathbf{x}) = \mathbf{P}_{01}(s,t;\mathbf{x})\mu(s+t;\mathbf{x}),$$

that can be sorted out as the following for differential purpose:

$$\mathbf{P}'_{00}(s,t;\mathbf{x}) + (\lambda(s+t;\mathbf{x}) + \mu(s+t;\mathbf{x}))\cdot\mathbf{P}_{00}(s,t;\mathbf{x}) = \mu(s+t;\mathbf{x}). \tag{A1}$$

Notice that

$$\frac{\partial}{\partial t}\left(\mathbf{P}_{00}(s,t;\mathbf{x})\cdot e^{\int_0^t \lambda(s+\tau;\mathbf{x})+\mu(s+\tau;\mathbf{x})d\tau}\right) = \mathbf{P}'_{00}(s,t;\mathbf{x})\cdot e^{\int_0^t \lambda(s+\tau;\mathbf{x})+\mu(s+\tau;\mathbf{x})d\tau} + \mathbf{P}_{00}(s,t;\mathbf{x})\cdot\frac{\partial}{\partial t}\left(e^{\int_0^t \lambda(s+\tau;\mathbf{x})+\mu(s+\tau;\mathbf{x})d\tau}\right).$$

Let $z \triangleq s+\tau$ and $dz \triangleq d\tau$. Then, the above equation is equivalent to:

$$\frac{\partial}{\partial t}\left(\mathbf{P}_{00}(s,t;\mathbf{x})\cdot e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\right) = \mathbf{P}'_{00}(s,t;\mathbf{x})\cdot e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz} + \mathbf{P}_{00}(s,t;\mathbf{x})\cdot\frac{\partial}{\partial t}\left(e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\right). \tag{A2}$$

Using the fact that:

$$\frac{\partial}{\partial t}\left(e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\right) = e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\cdot(\lambda(s+t;\mathbf{x}) + \mu(s+t;\mathbf{x})),$$

we may rewrite the Equation (A2) as the following:

$$\frac{\partial}{\partial t}\left(\mathbf{P}_{00}(s,t;\mathbf{x})\cdot e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\right) = e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}\cdot\left(\mathbf{P}'_{00}(s,t;\mathbf{x}) + (\lambda(s+t;\mathbf{x}) + \mu(s+t;\mathbf{x}))\cdot\mathbf{P}_{00}(s,t;\mathbf{x})\right). \tag{A3}$$

Combining Equations (A1) and (A3) and setting $S(s,t;\mathbf{x}) \triangleq e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}$ yields:

$$\frac{\partial}{\partial t}(S(s,t;\mathbf{x})\mathbf{P}_{00}(s,t;\mathbf{x})) = S(s,t;\mathbf{x})\left(\mathbf{P}'_{00}(s,t;\mathbf{x}) + (\lambda(s+t;\mathbf{x}) + \mu(s+t;\mathbf{x}))\mathbf{P}_{00}(s,t;\mathbf{x})\right) = S(s,t;\mathbf{x})\mu(s+t;\mathbf{x}).$$

Then, the general solution can be obtained by integrating both sides:

$$\int_0^t \frac{\partial}{\partial\xi}(S(s,\xi;\mathbf{x})\cdot\mathbf{P}_{00}(s,\xi;\mathbf{x}))\ \ d\xi = \int_0^t S(s,\xi;\mathbf{x})\cdot\mu(s+\xi;\mathbf{x})\ \ d\xi,$$

or equivalently:

$$S(s,t;\mathbf{x})\cdot\mathbf{P}_{00}(s,t;\mathbf{x}) = \int_0^t S(s,\xi;\mathbf{x})\cdot\mu(s+\xi;\mathbf{x})\ \ d\xi + C.$$

Substituting back $S(s,t;\mathbf{x})$ gives:

$$\mathbf{P}_{00}(s,t;\mathbf{x}) = \frac{1}{e^{\int_s^{s+t} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}}\left(\int_0^t e^{\int_s^{s+\xi} \lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ \ d\xi + C\right).$$

When $t=0$, we have $\mathbf{P}_{00}(s,0;\mathbf{x}) = 1$ for all $s \geq 0$ by definition, so that:

$$\mathbf{P}_{00}(s,0;\mathbf{x}) = \frac{1}{e^{\int_s^{s+0} \lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}}\left(\int_0^0 e^{\int_s^{s+\xi} \lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ \ d\xi + C\right) = \frac{1}{e^0}(0+C) = 1,$$

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

solving for the integral constant $C = 1$. Thus far, we have:

$$\mathbf{P}_{00}(s,t;\mathbf{x}) = \frac{1}{e^{\int_s^{s+t}\lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}}\left(\int_0^t e^{\int_s^{s+\xi}\lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ d\xi+1\right). \tag{A4}$$

Consider solving the denominator part of (A4):

$$e^{\int_s^{s+t}\lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz} = \exp\left(\int_s^{s+t}\frac{az^{a-1}}{1+z^a}\cdot e^{\mathbf{x}^\top\beta_\lambda}+\frac{bz^{b-1}}{1+z^b}\cdot e^{\mathbf{x}^\top\beta_\mu}\ dz\right).$$

Using the fact that:

$$\int_s^{s+t}\frac{az^{a-1}}{1+z^a}dz = \left[\log\left(1+z^a\right)\right]_s^{s+t} = \log\left(1+(s+t)^a\right)-\log\left(1+s^a\right),$$

we have:

$$\begin{aligned} e^{\int_s^{s+t}\lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz} &= \exp\left(\log\left(1+(s+t)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}+\log\left(1+(s+t)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}-\log\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}-\log\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}\right)\\ &= \frac{\left(1+(s+t)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+(s+t)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}. \end{aligned} \tag{A5}$$

Now, consider solving the numerator part of (A4):

$$\int_0^t e^{\int_s^{s+\xi}\lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ d\xi = \int_0^t\frac{\left(1+(s+\xi)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+(s+\xi)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}\cdot\frac{b(s+\xi)^{b-1}}{1+(s+\xi)^b}\cdot e^{\mathbf{x}^\top\beta_\mu}\ d\xi,$$

where the right-hand side can be sorted as the following:

$$e^{\mathbf{x}^\top\beta_\mu}\cdot\frac{1}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}\cdot\int_0^t\left(1+(s+\xi)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+(s+\xi)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}-1}b(s+\xi)^{b-1}\ d\xi.$$

Let $\theta \triangleq s+\xi$, $d\theta \triangleq d\xi$, and $u \triangleq 1+\theta^b$ so that $du \triangleq b\theta^{b-1}\ d\theta$ and $1+\theta^a \equiv 1+(u-1)^{a/b}$. Then,

$$\begin{aligned} \int_0^t e^{\int_s^{s+\xi}\lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ d\xi &= e^{\mathbf{x}^\top\beta_\mu}\cdot\frac{1}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}\cdot\int_s^{s+t}\left(1+\theta^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+\theta^b\right)^{e^{\mathbf{x}^\top\beta_\mu}-1}b\theta^{b-1}\ d\theta\\ &= e^{\mathbf{x}^\top\beta_\mu}\cdot\frac{1}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}\cdot\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\beta_\lambda}}u^{e^{\mathbf{x}^\top\beta_\mu}-1}\ du. \end{aligned} \tag{A6}$$

Combining Equations (A4)–(A6), we finally have the following transition probabilities:

$$\begin{aligned} \mathbf{P}_{00}(s,t;\mathbf{x}) &= \frac{1}{e^{\int_s^{s+t}\lambda(z;\mathbf{x})+\mu(z;\mathbf{x})dz}}\left(\int_0^t e^{\int_s^{s+\xi}\lambda(\gamma;\mathbf{x})+\mu(\gamma;\mathbf{x})d\gamma}\cdot\mu(s+\xi;\mathbf{x})\ d\xi+1\right)\\ &= \frac{e^{\mathbf{x}^\top\beta_\mu}\cdot\frac{1}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}\cdot\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\beta_\lambda}}u^{e^{\mathbf{x}^\top\beta_\mu}-1}\ du+1}{\frac{\left(1+(s+t)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+(s+t)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}{\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}}\\ &= \frac{e^{\mathbf{x}^\top\beta_\mu}\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\beta_\lambda}}u^{e^{\mathbf{x}^\top\beta_\mu}-1}\ du+\left(1+s^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+s^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}{\left(1+(s+t)^a\right)^{e^{\mathbf{x}^\top\beta_\lambda}}\left(1+(s+t)^b\right)^{e^{\mathbf{x}^\top\beta_\mu}}}, \end{aligned}$$

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and

$$\mathbf{P}_{01}(s,t;\mathbf{x}) = 1 - \mathbf{P}_{00}(s,t;\mathbf{x}) = 1 - \frac{e^{\mathbf{x}^\top\beta_\mu}\int_{1+s^b}^{1+(s+t)^b}\left(1+(u-1)^{a/b}\right)^{e^{\mathbf{x}^\top\beta_\lambda}}u^{e^{\mathbf{x}^\top\beta_\mu}-1}\,du + (1+s^a)^{e^{\mathbf{x}^\top\beta_\lambda}}(1+s^b)^{e^{\mathbf{x}^\top\beta_\mu}}}{(1+(s+t)^a)^{e^{\mathbf{x}^\top\beta_\lambda}}(1+(s+t)^b)^{e^{\mathbf{x}^\top\beta_\mu}}}.$$

For the cases that we do not consider the covariates so that $\mathbf{x} = 0$ and $e^{\mathbf{x}^\top\beta} = 1$, we have the following transition probabilities:

$$\begin{aligned}\mathbf{P}_{00}(s,t) &= \frac{\int_{1+s^b}^{1+(s+t)^b} 1+(u-1)^{a/b}\,du + (1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}\\ &= \frac{\left[\frac{b(u-1)^{a/b+1}}{a+b}+u\right]_{u=1+s^b}^{1+(s+t)^b} + (1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}\\ &= \frac{\frac{b\left((s+t)^{a+b}-s^{a+b}\right)}{a+b} + (s+t)^b - s^b + (1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)},\end{aligned}$$

and

$$\mathbf{P}_{01}(s,t) = 1 - \mathbf{P}_{00}(s,t) = 1 - \frac{\frac{b\left((s+t)^{a+b}-s^{a+b}\right)}{a+b} + (s+t)^b - s^b + (1+s^a)(1+s^b)}{(1+(s+t)^a)(1+(s+t)^b)}.$$

The derivation procedures for $\mathbf{P}_{11}(s,t;\mathbf{x})$, $\mathbf{P}_{10}(s,t;\mathbf{x})$, $\mathbf{P}_{11}(s,t)$, and $\mathbf{P}_{10}(s,t)$ are essentially the same with the switched rate functions.

A.2 R code for simulations

The R code for NH-CTMC simulations is on the following GitHub repository: https://github.com/jchang49/NH-CTMC.

10970258, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sim.9707 by University Of Texas - Ham/Tmc, Wiley Online Library on [21/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License