Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Modelling transitions of opioid usage, addiction, and fatal overdoses using a Bayesian multistate model

Michael Jongho Moon* | Monica Alexander

[1]Department of Statistical Sciences, University of Toronto, Ontario, Canada

Correspondence
*Michael Jongho Moon, Department of Statistical Sciences, University of Toronto, Ontario, Canada. Email: moon@utstat.toronto.edu

Present Address
Department of Statistical Sciences, University of Toronto, Ontario, Canada

Summary

Opioid addiction and overdoses are an ongoing crisis in North America. During the 1990s and 2000s, increases in overdoses were mostly observed in the non-Hispanic population. However, in more recent years, the proliferation of synthetic opioids, such as fentanyl, has contributed to substantial death toll increases particularly in the non-Hispanic population. While there has been much previous research on understanding trends of opioid-related deaths, there has been less focus on understanding disparities and trends at opioid-related stages before death, namely opioid use, misuse, and non-fatal overdose and treatment. We propose a theoretical framework in understanding the transitions that lead to illicit opioid use and overdose deaths using a Bayesian multistate approach. Our framework's multistate model divides opioid use stages into prescription opioid use, opioid misuse, addiction treatment, and opioid-induced death. Bayesian back-calculation allows us to estimate time-variant transition probabilities and incidences. Simulation studies demonstrated that our framework can retract the transition probabilities with informative priors while its incidence estimates may be imprecise. We then applied the proposed framework to non-Hispanic black and non-Hispanic white populations in the United States from 2015 to 2019 integrating multiple sources of public data. We analyzed the posterior estimates of the transition probabilities to study the disparities in opioid use between the two groups. The analysis result suggests that the non-Hispanic black population faced higher, and considerably increasing, risk of opioid-induced deaths conditional on using opioids in an illicit manner compared to the non-Hispanic white population.

KEYWORDS:
opioid, mortality, multistate modelling, Bayesian, back-calculation

1 | INTRODUCTION

Opioid addiction and overdoses are an ongoing crisis in North America. In 2020, 68 630 people died from drug overdose involving opioids in the United States[1], a 38% increase from the previous year. During the same period 6 415 people died in Canada from opioid-related overses[2], a 73% increase from the previous year. Historically, there have been marked racial disparities in the opioid epidemic. During the 1990s and early 2000s, increases in overdoses were mostly observed in the non-Hispanic white

population.[3] However, in more recent years, the proliferation of synthetic opioids, such as fentanyl, has contributed to the rising death toll rates for all populations, with particularly substantial increases in the non-Hispanic black population, with other important differences by age and geographic region.[4]

While there has been much previous research on understanding trends of opioid-related deaths by key demographic subgroups[3,5,6], there has been less focus on understanding disparities and trends at opioid-related stages before death, namely opioid use, misuse, and non-fatal overdose and treatment. Some work has incorporated the intermediate stages. For example, Pitt et al.[7] used a dynamic compartmental model to describe the transitions between different stages divided based on prescription opioid use and addiction, treatment, and levels of pain. Battista et al.[8] suggested a compartmental model that described the dynamics between prescription opioid users and those suffering from opioid addiction. However, in both cases, they used individually estimated transition rates from external sources. To date there is not a clear unified picture of the rate at which individuals transition from prescription to misuse, then potentially from misuse to overdose, and how this differs across racial and ethnic lines. Such information would be valuable in designing targeted interventions to help curb opioid addiction and prevent loss of life.

The challenge with estimating such transition rates is one of data availability. While good quality cause-of-death data generally exist in the United States and Canada, the data available to understand population-level incidence of opioid use, and opioid misuse, is not widely available. Partial information may be available from surveys, but these data have large levels of uncertainty compared to data from vital registration systems.

In this paper, we propose a theoretical framework in understanding the transitions that lead to illicit opioid use and overdose deaths using a Bayesian multistate modelling approach. The model utilizes different sources of data on opioid use and deaths in a single framework, and allows reliable estimates of all transition rates. The method builds on work by Brookmeyer and Gail[9] who used a back-calculation technique to estimate unobserved HIV infection rates from observed AIDS diagnosis rates. Sweeting et al.[10] then proposed a Bayesian inference of the technique. Our framework provides probabilistic estimates for the transition probabilities in a multistate model that captures both prescription-initiated and illicitly-initiated opioid users. The Bayesian inference incorporates partial observations of interim transitions to augment data on opioid-induced mortality.

In the following section, we first survey the existing literature on estimating prevalence and incidence of illicit opioid use as well as those that study transitions in multistate models in other contexts. The Method section then introduces the proposed framework in detail and the Simulation section presents simulation studies to investigate the feasibility and robustness of the proposed framework. The Case study section follows where we apply the method to the opioid crisis in the United States to study the racial disparities in the transition behaviour. Finally, we conclude the paper with key implications and limitations of the proposed framework in the Discussion.

2 | RELATED WORK

2.1 | Compartmental modelling of drug addiction dynamics

Quantifying the occurrence rate of new cases (i.e., incidence) and the frequency of current cases (i.e., prevalence) are important goals for understanding the spread of drug addiction[11]. To quantify the incidence and prevalence of opioid addiction, we may utilize compartmental models which are popular in epidemiology for modelling infectious diseases. We find such models used for opioid addiction in the literature where each drug user is analogous to a contagion that spread the innovation of the drug[12,13,14,15]. They provide a conceptual framework useful for studying the epidemiology of opioid addiction and overdose deaths.

In epidemiology, compartmental models describe population classes in the epidemic of an infectious disease.[16] For example, classic SIR models divide the population into the susceptible (S), the infective (I), and the recovered (R). The adoption of epidemiological models for opioid addiction started as early as 1979 when Douglas and Stewart[13] adopted a simplified SIR model to study the spread of heroin use. The simplified model described the heroin usage progression with three compartments - the susceptible, the affected including both drug users and non-users involved in the proliferation of the drug, and those removed from the affected group for any reason. In 2007, White and Comiskey[15] introduced a fully adopted SIR model for the global heroin epidemic. The White-Comiskey model described the susceptible, the drug users who are not in treatment, and the drug users who are in treatment. Furthermore, the model depicted inflows to the susceptible group and outflows from each of the three states. The outflows included both natural and drug-related deaths as well as other types of removals. Figure 1 displays the compartments and the transitions defined by White and Comiskey.[15]

There have been efforts to extend the White-Comiskey model to capture a more realistic and detailed dynamics of the spread of opioid use. Rossi[17] adopted a Mover-Stayer model from HIV/AIDS epidemic modelling to dividing the susceptible population

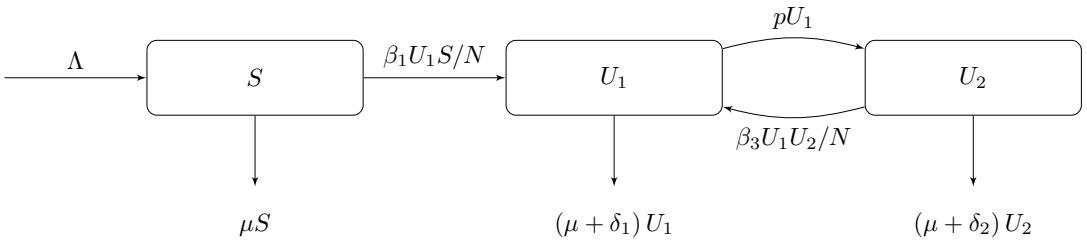


Figure 1 An opiate-using career model by White and Comiskey[15]. S, U_1, and U_2 denote the susceptible population, drug users not in treatment, and drug users in treatment respectively. Λ is the number of individuals entering the susceptible population, N the total population, μ the natural death rate, and δ_1 and δ_2 drug-related removal rates. β_1, p, and β_2 represent the probabilities for becoming a drug user, entering treatment, and relapsing to untreated drug use from treatment respectively.

into the Stayers — those that are not at risk due to their "prudent" behaviour — and the Movers who are at risk. The proposed model also added states to describe the latency period during which an individual's drug usage unobserved. The model consists of six compartments excluding the general population and deaths. On the other hand, Djilali et al.[18] generalized the White-Comiskey model by introducing a non-linear incidence and treat-age affect for the relapse probability. The treat-age affect captures the behavioural change of the drug users under treatment. The presented non-linear incidence function depends on both the susceptible population size (S) and the number of drug users not in treatment (U_1). The authors present the conditions for the non-linear incidence function that allow asymptotic stability of the drug-free equilibrium and asymptotic stability of the endemic equilibrium.

These conceptual models capture the size of different epidemiological classes, or compartments, of the population at a given time (prevalence) and the rate of transition between these compartments (incidence). They use mathematical expressions to relate the prevalence and incidence. In the following subsections, we discuss how back-calculation techniques are used to provide empirical estimates for the prevalence and incidence.

2.2 | Estimation of prevalence and incidence via back-calculation

Back-calculation techniques allow empirical estimation of parameters within opioid incidence and prevalence models. Brookmeyer and Gail[9] first proposed a back-calculation technique to estimate unobserved HIV incidences from AIDS diagnosis counts. The technique has since been adopted to estimate the unobserved drug misuse incidence and prevalence.[19,20] The methods define the dynamics of a contagious innovation's proliferation with a convolution relationship between the incidence, the transition distribution, and the endpoint distribution. They assume that observed counts at an endpoint follow a probabilistic distribution defined by the convolution relationship and trace back the relationship to estimate the incidence and prevalence.

de Angelis et al.[19] used opiate-related deaths recorded in England between 1968 and 2000 to estimate new opiate injections. The back-calculation method assumed fixed mortality rates that are age-specific and fixed cessation rates per year. These values served as the endpoint counts in their study. In the case of the cessation rates, the study used three different estimates to account for uncertainties and differences from different sources. The study modeled drug abusers as a single state and estimated the age-specific incidences using a maximum likelihood algorithm.

A study by Sánchez-Niubò et al.[20] similarly adopted the back-calculation method to estimate the unobserved heroin incidences in Spain between 1971 and 2005. Figure 2 reconstructs the proposed multistate model diagram from the paper. The study based its estimates on the incidence of new opioid treatments instead of mortality data. The authors used a mix of observed data as well as independent estimates for interim transition rates (p_t and q_t). These values were, in turn, used to estimate h_t by maximizing the Poisson likelihoods of the records of new opioid treatment admissions.

The back-calculation methods allow study of long-term trends in incidence and prevalence of opioid misuse. They are particularly useful in the context of opioid misuse and overdose as there is limited data availability on incidence and prevalence due as they are difficult to measure directly and highly biased when reported.[19] On the other hand, the maximum likelihood approaches do not provide direct estimation of the uncertainties associated with the estimates. Instead, both groups[19,20] constructed confidence intervals based on resamples generated using bootstrap techniques.

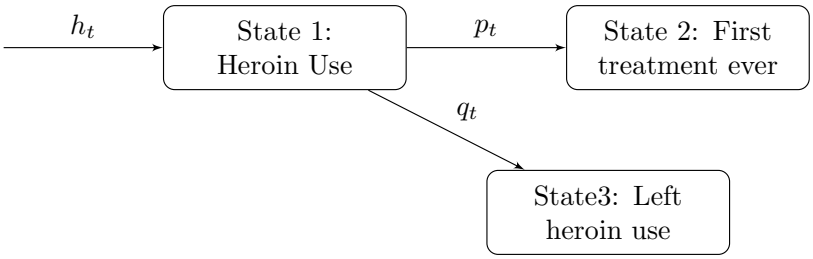


Figure 2 Multistate model diagram by Sánchez-Niubò et al.[20]. Parameter h_t denotes the expected number of people starting heroin use, p_t transition rate entering treatment for the first time, and q_t transition rate leaving heroin use without entering treatment, at time t.

2.3 | Extending back-calculation with the Bayesian approach

Aalen et al.[21] extended the original back-calculation method by Brookmeyer and Gail[9] to incorporate a multistate Markov modelling of HIV progression. Since then, Bayesian back-calculation methods have emerged extending the original model further. The Bayesian approach allow probabilistic estimation of each parameter of interest and provide the level of uncertainty associated with the parameter estimate.

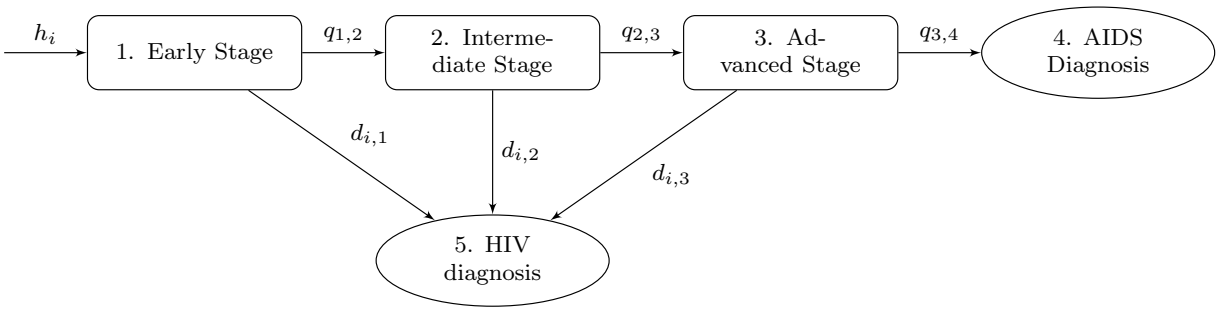


Figure 3 A multistate model for HIV progression.[10] h_i represents new HIV infections during i^{th} time interval. $q_{j,k}$ represents HIV progression rates among undiagnosed HIV carriers from j^{th} stage to k^{th} stage based on CD4 counts. The model assumes undiagnosed HIV carriers at j^{th} stage are diagnosed with HIV with a probability of $d_{i,j}$ during i^{th} time interval.

Sweeting et al.[10] first proposed a Bayesian multistate model for back-calculation to estimate the HIV incidences. Figure 3 shows the multistate HIV progression model used. The Bayesian approach provided posterior distribution estimates for the incidences and HIV diagnosis probabilities from observed AIDS and HIV diagnoses counts. Their simulation study demonstrated that the Bayesian provided reasonable posterior estimates that traced the true parameter values. The authors also demonstrated incorporating additional data on HIV diagnoses at different stages to further improved the estimates. Birrell et al.[22] adopted the method to investigate other key epidemiological quantities for an updated data of the same epidemic. They showed that the method can inform distributions of the time-to-diagnosis, the time-since infection, and the prevalence of undiagnosed infection.

Brizzi et al.[23] further extended the Bayesian multistate model for back-calculation to estimate age and time dependent HIV incidences. The extended method modeled the infection process as a two-dimensional non-homogeneous Poisson process with both age and time dependent infection rates. The method no longer assumes a first-order Markov chain, but assumes that progression and diagnosis probabilities depend on age and time at infection. While the additional dynamics substantially complicated the model, the authors reported reasonably well-fitting posterior estimates using bivariate splines for the log-incidence surface.[23]

Extending the back-calculation method with a Bayesian multistate model provided model-based estimation of the uncertainties via posterior distributions while incorporating additional data into the model. In this study, we adopt and extend the Bayesian multistate back-calculation in the context of the drug use and mortality.

3 | METHOD

3.1 | A multistate model for opioid use and mortality

We propose a discrete-time multistate model to describe stages of opioid use as shown in Figure 4. The model depicts the transitions in discrete time intervals $\left(t_{i-1}, t_i\right]$ for $i = 1, \ldots, T$ and incorporates two independent entry points to distinguish those those who have used prescription opioids before misusing opioids from those who initiate illicit opioid use directly. $r_i^{(1)}$ denotes the expected rate of new prescription opioid users entering State (1) during $\left(t_{i-1}, t_i\right]$. $r_i^{(2)}$ denotes the expected rate of new illicit opioid users entering State (2) in $\left(t_{i-1}, t_i\right]$. Both entries represent occurrence rates of opioid users (i.e., incidences) among those who never used opioid prior to entering the model.

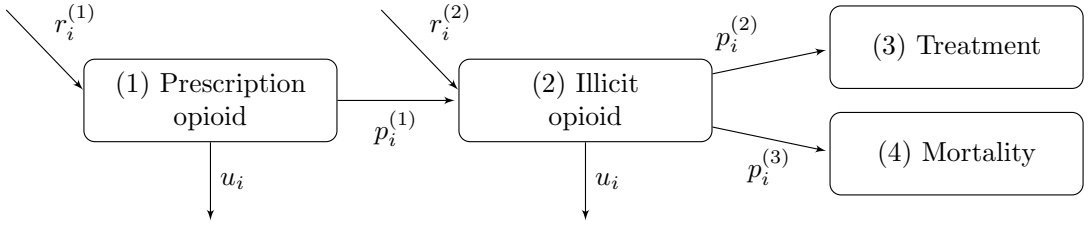


Figure 4 Transition diagram of the multistate model. $r_i^{(1)}$ and $r_i^{(2)}$ represent the incidences for prescription and illicit opioid users respectively during $\left(t_{i-1}, t_i\right]$. $p_i^{(1)}$, $p_i^{(2)}$, and $p_i^{(3)}$ denote the transition probabilities between the states indicated by the arrows during $\left(t_{i-1}, t_i\right]$. u_i denotes the non-opioid-related mortality rate.

$p_i^{(m)}$ for $m = 1, 2, 3$ denote the transition probabilities between different states in the model during $\left(t_{i-1}, t_i\right]$. $p_i^{(1)}$ represents the probability of misusing opioids among those who have used prescription opioids. The probability of an illicit opioid user entering treatment, State (3), is denoted by $p_i^{(2)}$. Both $p_i^{(1)}$ and $p_i^{(2)}$ are conditional on not having deceased due to reasons unrelated to opioid overdose by t_{i+1}. $p_i^{(3)}$ represents the probability of experiencing an overdose death, State (4), among those who have misused opioids by $\left(t_{i-1}, t_i\right]$. The probability is conditional on not having entered treatment in addition to not having deceased due to other reasons. u_i represents the mortality rate that are unrelated to opioid overdose.

We emphasize that our proposed model aims to provide an aggregate depiction and that it focuses on a subset of all possible transitions. For example, we assume that State (4) is an absorbing state. However, those who receive treatment may start using illicit opioids again. Similiarly, the model also does not distinguish those who stopped using opioids without entering a treatment facility from those who are actively using opioids. The model simplifies the real-world with these omissions but captures the main transitions of interest.

3.2 | The back-calculation of the incidences and transition probabilities

Back-calculation techniques allow estimating unobserved occurrence rate of new cases based on observed endpoint counts in a multistate model. The convolution relationship between the incidences, transition distribution, and endpoint counts is central to back-calculation. For example, Equation 1 describes a homogeneous convolution relationship in a discrete-time setting with disjoint sub-intervals, $\left(t_{i-1}, t_i\right]$ for $i = 1, \ldots, T$.

$$\mu_i = \sum_{\ell=1}^{i} h_\ell f_{\ell,i}, \tag{1}$$

where μ_i is the expected number of the endpoint counts in the time interval $\left(t_{i-1}, t_i\right]$, h_ℓ is the expected incidences in the interval $\left(t_{\ell-1}, t_\ell\right]$, and $f_{\ell,i}$ is the probability that an individual infected in $\left(t_{\ell-1} j, t_\ell\right]$ arrives at the endpoint in $\left(t_{i-1}, t_i\right]$, for $i = 1, \ldots, T$ and $\ell \leq i$. The back-calculation technique allows reconstruction of the expected incidences based on the observed number of endpoint arrivals and transition distributions. For example, the method is used to estimate of the incidences of HIV infections, h_ℓ, from the observed clinical diagnosis of AIDS, μ_i, and the incubation distribution that describe the progression from infection to AIDS, $f_{\ell,i}$.[10]

We introduce a back-calculation technique that is tailored to estimating the transition distributions as well as the incidences for opioid users. Equation 2 and Equation 3 describe the dynamics of the proposed multistate model. λ_i is a 4×4 matrix whose

(m, n)th entry represent the transition probability from State (m) to State (n) during $\left(t_{i-1}, t_i\right]$. Equation 2 defines the probabilities.

$$\left(\lambda_i\right)_{m,n} = \begin{cases} \left(1-u_i\right)\left(1-p_i^{(1)}\right) & m = n = 1 \\ \left(1-u_i\right)\left(1-p_i^{(2)}\right)\left(1-p_i^{(3)}\right) & m = n = 2 \\ \left(1-u_i\right)p_i^{(m)} & m \in \{1,2\}, n = m+1 \\ \left(1-u_i\right)\left(1-p_i^{(2)}\right)p_i^{(3)} & m = 2, n = 4 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

For example, a prescription opioid user in State (1) begins to use opioids in a manner that is not directed by their doctor with a probability of $\left(\lambda_i\right)_{1,2} = \left(1-u_i\right)p_i^{(1)}$ during $\left(t_{i-1}, t_i\right]$. The same person is then at risk of a opioid-induced death during $\left(t_i, t_{i+1}\right]$ with a probability of $\left(\lambda_i\right)_{2,4} = \left(1-u_{i+1}\right)\left(1-p_{i+1}^{(2)}\right)p_{i+1}^{(3)}$.

$$\boldsymbol{\Lambda}_i = \lambda_i^T \boldsymbol{\Lambda}_{i-1} + \left(r_i^{(1)}, r_i^{(2)}, 0, 0\right)^T \tag{3}$$

for $i = 1, \ldots, T$.

3.3 | Bayesian inference of transition probabilities and incidences

We use Bayesian inference to estimate the transition probabilities and the incidences. We assume that we have reliable yearly data on opioid-induced mortality counts, opioid addiction treatment initiate counts, and opioid misuse initiate counts. We also assume the non-opioid-induced mortality rates are known. Let x_i, y_i, an w_i be the observed counts for new opioid-induced deaths, treatment initiates, and misuse initiates for each $\left(t_{i-1}, t_i\right]$. We assume such data are readily available and demonstrate a real-world application using multiple data sources in Section 5.

To simplify notations, we let $\boldsymbol{r}_i = \left(r_i^{(1)}, r_i^{(2)}\right)^T$ and $\boldsymbol{p}_i = \left(p_i^{(1)}, p_i^{(2)}, p_i^{(3)}\right)^T$. We model the observed counts using Poisson distributions as shown in Equations 4, 5, and 6 with the means defined using the back-calculation model.

$$\begin{aligned} \left(x_i \left| \boldsymbol{\Lambda}_{i-1}, \boldsymbol{r}_i, \boldsymbol{p}_i, u_i \right.\right) &\sim \text{Poisson}\left(X_i\right) \\ \text{where} \quad X_i &= \Lambda_{i,4} \quad \text{for} \quad i = 1, \ldots, T \end{aligned} \tag{4}$$

$$\begin{aligned} \left(y_i \left| \boldsymbol{\Lambda}_{i-1}, \boldsymbol{r}_i, \boldsymbol{p}_i, u_i \right.\right) &\sim \text{Poisson}\left(Y_i\right) \\ \text{where} \quad Y_i &= \Lambda_{i,3} \quad \text{for} \quad i = 1, \ldots, T \end{aligned} \tag{5}$$

$$\begin{aligned} \left(w_i \left| \Lambda_{i-1}, \boldsymbol{r}_i, \boldsymbol{p}_i, u_i \right.\right) &\sim \text{Poisson}\left(W_i\right) \\ \text{where} \quad W_i &= \Lambda_{i-1,1}\left(1-u_i\right)p_i^{(2)} + r^{(2)_i} \quad \text{for} \quad i = 1, \ldots, T \end{aligned} \tag{6}$$

We use a hierarchical Bayesian framework to specify the expected incidences $\boldsymbol{r}_i$ and transition probabilities $\boldsymbol{p}_i$. To allow smooth curves over time, we model $\gamma_i^{(\cdot)} = \log\left(r_i^{(\cdot)}\right)$ and $\delta_i^{(\cdot)} = \text{logit}\left(p_i^{(\cdot)}\right)$ as Gaussian random walk processes. Equation 7 and Equation 8 describe the random walk models.

$$\gamma_i^{(m)} \left| \gamma_{i-1}^{(m)} \right. \sim N\left(\gamma_{i-1}^{(m)}, \sigma_\gamma^2\right) \tag{7}$$

$$\delta_i^{(n)} \left| \delta_{i-1}^{(n)} \right. \sim N\left(\delta_{i-1}^{(n)}, \sigma_\delta^2\right) \tag{8}$$

for $m = 1, 2$, $n = 1, 2, 3$, and $i = 2, 3, \ldots, T$. We specify the prior distributions of standard deviations of the Gaussian random walks 7 and 8, σ_γ and σ_δ, using half-normal distributions with standard deviations ν_γ and ν_δ respectively. The hyperparameters $\boldsymbol{r}_0$, $\boldsymbol{p}_0$, ν_γ, and ν_δ are specified as part of a Bayesian analysis. The initial prevalence of opioid prescription users, $\Lambda_{0,1}$, and those who have misused opioids, $\Lambda_{0,2}$, are also specified as part of a Bayesian analysis. For convenience, we specify $\Lambda_{0,3} = \Lambda_{0,4} = 0$.

3.4 | Incorporating partial data on transitions

In addition to the count data, we extend the model by assuming that we partially observe the transitions over the multistate model. This allows us to incorporate data on transitions that are observed on a subset of the population of interest. For example, the U.S. National Survey on Health and Drug Use[24] captures information on heroin use, prescription opioid misuse, and addiction treatments. The survey has included questions on regular use of prescription opioids since 2015.[24] From the survey responses,

we can infer the proportions of prescription opioid users who started using illicit opioids and illicit opioid user who entered a treatment facility in each year.

We denote the partially observed counts at t_{i-1} of those on prescription opioids as $S_{i-1}^{(1)}$ and of those misusing opioids as $S_{i-1}^{(2)}$. We let $s_i^{(1)}$ and $s_i^{(2)}$ denote the number of observed transitions to illicit opioid use out of those in $S_{i-1}^{(1)}$ and the number of observed transitions to treatment out of those in $S_{i-1}^{(2)}$, respectively by t_i. Assuming representativeness of the observations, we incorporate the observed partial transitions with the normal approximation described in Equation 9.

$$
\begin{aligned}
p_i^{(m),\text{part}} &\sim N\left(p_i^{(m)} + \phi^{(m)}, \eta_{i-1}^{(m)}\right) \\
\text{where} \quad p_i^{(m),\text{part}} &= \frac{s_i^{(m)}}{S_{i-1}^{(m)}} \quad \text{and} \quad \eta_{i-1}^{(m)} = \frac{p_i^{(m)}\left(1 - p_i^{(m)}\right)}{S_{i-1}^{(m)}} \\
&\text{for} \quad m = 1, 2 \quad \text{and} \quad i = 2, 3, \ldots, T
\end{aligned}
\tag{9}
$$

$\phi^{(1)}$ and $\phi^{(2)}$ allows any bias introduced by survey estimates. There may also be systematic bias introduced during the administration of surveys. Even when surveys guarantee unbiased estimates, the proportion of two estimators does not guarantee an unbiased estimator.

We specify the priors for $\phi^{(1)}$ and $\phi^{(2)}$ as part of a Bayesian analysis. If we have some prior knowledge about the likely direction of systematic bias of introduced during data collection, this information can be encoded into the model through the prior. For example, if we believe the survey under-reports illicit opioid use, we may specify the priors for $\phi^{(\cdot)}$ to follow a truncated normal distribution between -1 and 1 with a negative mean value.

4 | SIMULATION

In this section, we investigate the feasibility and robustness of the proposed methods described in Section 3 using a set of simulated data. Table 1 lists the 'true' parameters used for the simulation. $\log\left(r_0^{\text{TRUE},(m)}\right)$ and $\text{logit}\left(p_0^{\text{TRUE},(m)}\right)$ were used as initial means in the Gaussian random walks for log-incidences and logit-transition probabilities, respectively. The random walks simulated incidences and transition probabilities with standard deviations $\sigma_\gamma^{\text{TRUE}}$ and $\sigma_\delta^{\text{TRUE}}$ for time periods $\left(t_{i-1}, t_i\right]$, $i = 0, \ldots, 20$. We then used the simulated incidences and probabilities to simulate opioid-use progression data following the multistate model defined by Equation 2. We assumed a time-invariant non-opioid-induced mortality rate per time period, u^{TRUE}. The counts of opioid-induced deaths, x_i, treatment initiates, y_i, and opioid misuse initiates, w_i were recorded to simulate observed data for $i = 1, \ldots, 20$. Additionally, we simulated a survey where a random sub-sample of the simulated individuals were captured. We sampled 5% among those using prescription opioids and those using illicit opioids at each time period. All codes used for the simulation studies are available at https://github.com/mjmoon/bmm-opioid.

4.1 | Feasibility of the method

We investigated the feasibility of the method by fitting models using priors that mimic the 'true' parameters used to simulate the data. The parameter values used for initial values were the same values as the simulation parameters - $\boldsymbol{r}_0 = \left(r_0^{\text{TRUE},(1)}, r_0^{\text{TRUE},(2)}\right)$, $\boldsymbol{p}_0 = \left(p_0^{\text{TRUE},(1)}, p_0^{\text{TRUE},(2)}, p_0^{\text{TRUE},(3)}\right)$, $\sigma_{\gamma_0} = \sigma_\gamma^{\text{TRUE}}$, and $\sigma_{\delta_0} = \sigma_\delta^{\text{TRUE}}$. For the hyperparameters of the random walks, we use non-informative half-normal priors with $\nu_\gamma = 1$ and $\nu_\delta = 1$. For $\phi^{(1)}$ and $\phi^{(2)}$, we utilized non-informative priors using standard normal distributions truncated on $(-1, 1)$. We used RStan's[25,26] implementation of No-U-Turn sampler[27] to estimate the posterior distributions with and without the partial data on transitions from the simulated survey data.

Figure 5 displays the estimation results for the transition probabilities and the incidences. The left column is the result from incorporating the ratio estimates, $p_i^{(1),\text{part}}$ and $p_i^{(2),\text{part}}$, from the simulated survey data, and the right column the result without the ratio estimates.

Incorporating the ratio estimates was the most beneficial in estimating the transition probabilities from prescription to illicit use of opioids, $p_i^{(1)}$. Without the ratio estimates, the posterior medians remained relatively constant over time. The 95% credible intervals also missed 15 out of the 20 simulation parameters. On the other hand, the posterior distributions that incorporated the ratio estimates successfully traced the simulation parameters of the target values over time. The 95% credible intervals

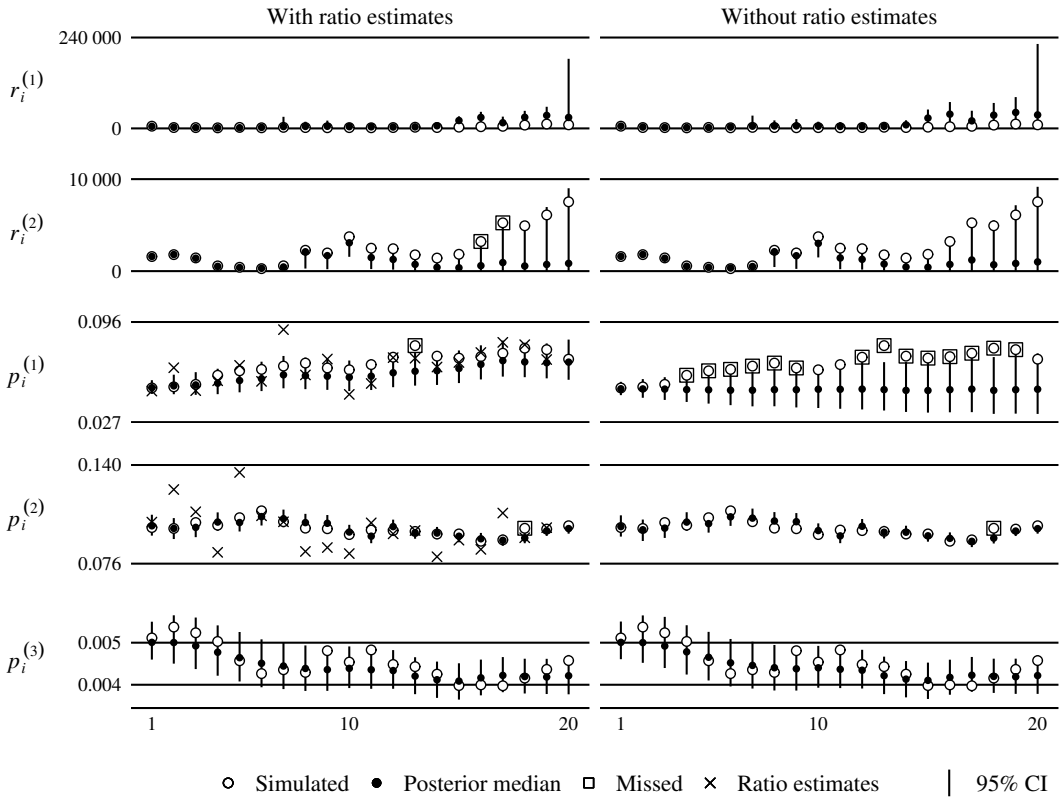


Figure 5 Simulated true values and posterior medians with 95% credible intervals for the incidences and transition probabilities using the baseline priors. The left column displays results from the analysis that incorporates the ratio estimates based on simulated surveys. Without the ratio estimates, the model estimates for $p_i^{(1)}$ remain centred around the initial value over time and the credible intervals miss the true values.

were smaller and contained the simulation parameters for all but 1 time period. The posterior distributions for other transition probabilities and incidences show very little difference between the two columns in Figure 5.

The result also shows a better performance at estimating the transition probabilities in comparison to estimating the incidences. Regardless the use of ratio estimates, the credible intervals for the incidences increased exponentially with time. In both cases, the posterior medians consistently overestimated the opioid prescription incidences, $r_i^{(1)}$, while consistently underestimating the direct opioid misuse incidences, $r_i^{(2)}$ for time periods with $i = 7$ and beyond. In contrast, the intervals for the transition probabilities remained relatively stable.

The feasibility study reveals both optimistic results and limitations of the proposed framework. With the ratio estimates incorporated, the framework was capable of making reliable inference about the transition probabilities. For all transition probabilities, the posterior medians traced the parameters over time. It is also notable that the posterior distributions remained consistently closer to the parameters than the ratio estimates. The inferences on the incidences were not as reliable over time. The posterior distributions grew exponentially with time making the inference about them ambiguous for the latter half of the periods. While the framework may provide reasonable estimates for the short periods, we emphasize that our focus is on the transition probabilities.

4.2 | Sensitivity to prior choices

We conducted a sensitivity analysis to investigate the proposed framework's robustness to prior choices. We considered the model with ratio estimates from Section 4.1 as the reference model, and fitted models with a subset of priors deviated from the reference model at different levels each time. We considered the following cases using $\alpha \in \{0.5, 2\}$.

(SA1) $\nu_\gamma = \alpha \cdot \nu_\gamma^*$ where ν_γ^* is the ν_γ value used in the reference model

(SA2) $\nu_\delta = \alpha \cdot \nu_\delta^*$ where ν_δ^* is the ν_δ value used in the reference model

(SA3) $\boldsymbol{r}_0 = \alpha \cdot \left(r_0^{\text{TRUE},(1)}, r_0^{\text{TRUE},(2)}\right)$

(SA4) $\boldsymbol{p}_0 = \alpha \cdot \left(p_0^{\text{TRUE},(1)}, p_0^{\text{TRUE},(2)}, p_0^{\text{TRUE},(3)}\right)$

In all scenarios, we fitted models incorporating the ratio estimates and using the same algorithm as the reference model.

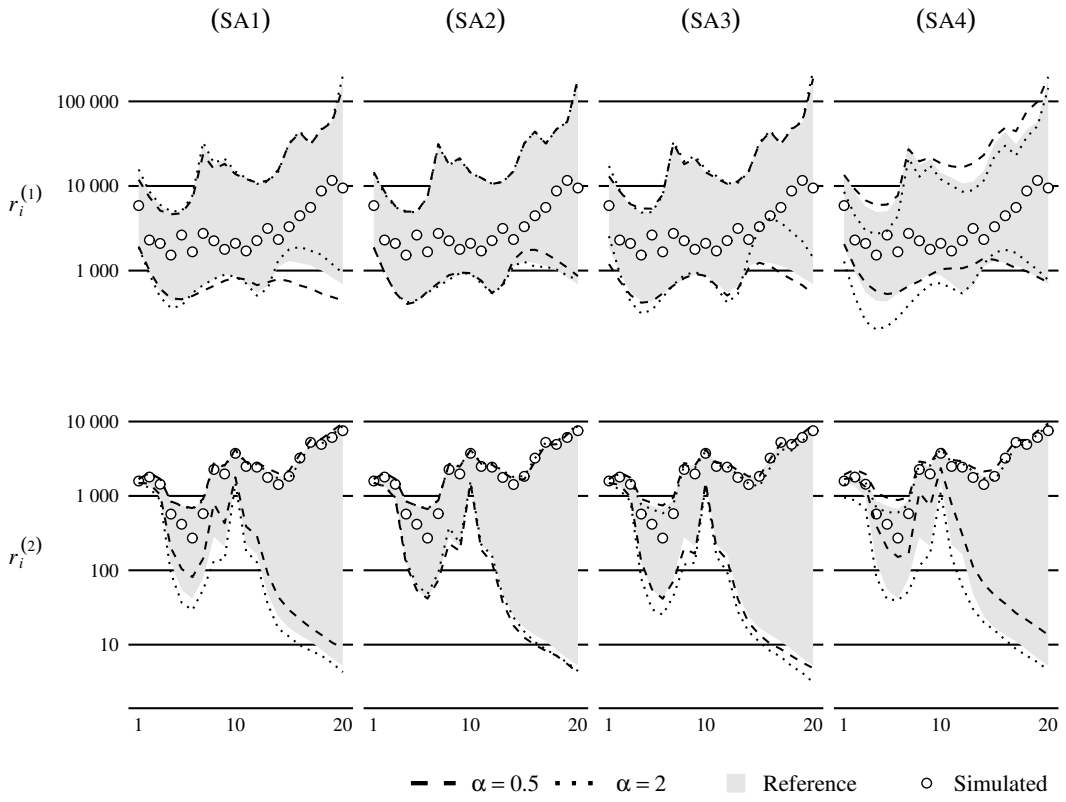


Figure 6 95% credible intervals of the incidences from the sensitivity analysis cases. For each case, we considered doubling the priors, $\alpha = 2$, and halving, $\alpha = 0.5$. The intervals are plotted in log-scale to highlight the deviations. In all cases, the intervals contain the simulated parameters despite deviations from the reference model's intervals.

Figure 6 displays the effects of the deviations in priors on the incidence posterior distributions. The dotted lines and dashed lines represent the 95% credible intervals from sensitivity analysis models with $\alpha = 2$ and $\alpha = 2$ respectively. The reference model's credible intervals are shown in gray shades and the simulated parameters as points for comparison. Using different values of ν_γ in (SA1) and different $\boldsymbol{r}_0$ values in (SA3) caused minor deviations in the credible intervals. While these priors directly concern the incidences over time, the prior choices did not cause significant differences in the posterior distributions. Using different values of ν_δ which concerns how transition probabilities change over time in (SA2) made even less difference on the incidence estimates. In contrast, changing the initial transition probabilities, $\boldsymbol{p}_0$, in (SA4) resulted in the largest amount of deviations from the reference model. However, the overall trend remained consistent with the reference model.

Figure 7 shows of the transition probability posterior distributions resulting from the sensitivity analysis. Changing ν_γ in (SA1), ν_δ in (SA2), and $\boldsymbol{r}_0$ in (SA3) did not result in any visible deviations in the posterior distributions from those of the reference model. Changing the initial transition probabilities, $\boldsymbol{p}_0$, in contrast, had clear effects on the posterior distributions. In particular, the effects of shifting the priors sustained throughout the 20 time periods for the transition probabilities from opioids

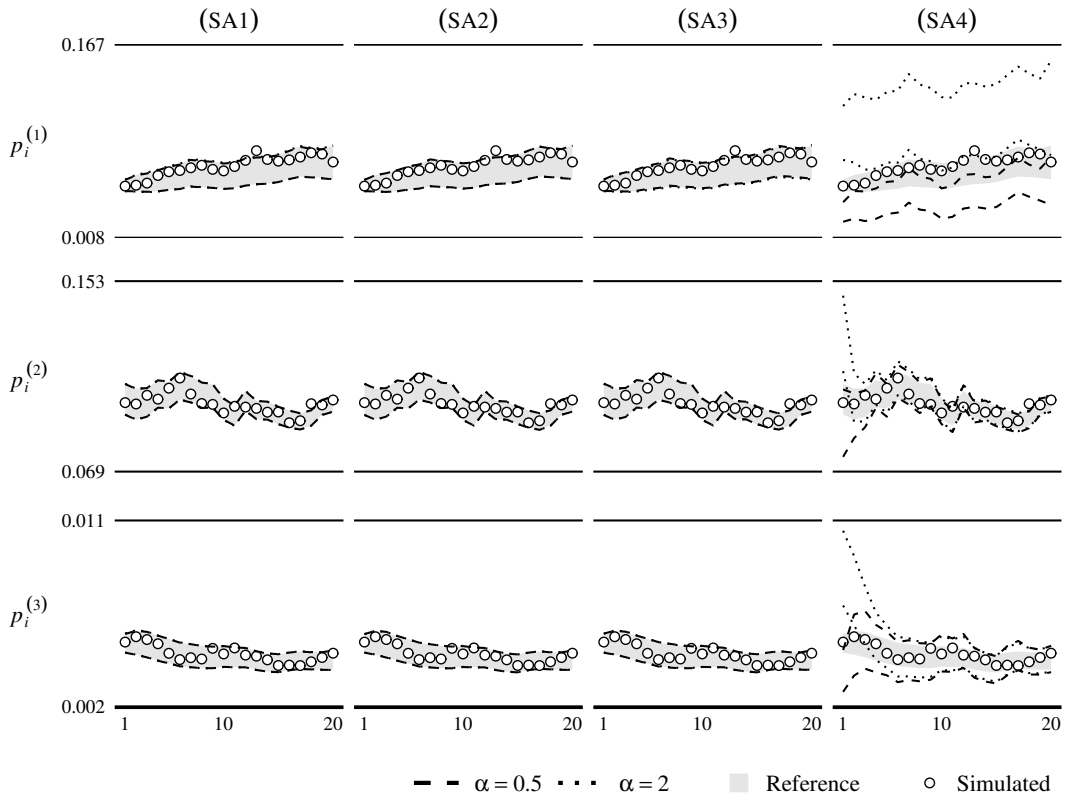


Figure 7 95% credible intervals of the transition probabilities from the sensitivity analysis cases. For each case, we considered doubling the priors, $\alpha = 2$, and halving, $\alpha = 0.5$. While changing other priors did not make any visible difference in the posterior distributions, updating the initial transition probabilities, (SA4), resulted in visible deviations from the reference model's intervals. In particular, the credible intervals for $p_i^{(1)}$ from (SA4) consistently miss the simulated parameters.

prescription to misuse, $p_i^{(1)}$. The posterior distributions of $p_i^{(2)}$ and $p_i^{(3)}$ displays the effects of the shifts in the initial values but the effects do not last long. They were consistent with the reference posterior distributions within 2 to 5 time periods.

The proposed framework was reasonably robust to prior choices for the random walk variances and the initial values for the incidences. However, it is sensitive to the choice of the initial transition probabilities. This suggests that the framework would benefit from informative priors for the initial transition probabilities. In the case where informative priors are not available, the inference about the transition probabilities over time must should be done with caution. In particular, the posterior distributions of $p_i^{(1)}$ maybe biased over all time periods.

5 | CASE STUDY: THE OPIOID USE IN THE U.S.

We demonstrate a real-world application of our proposed model to the opioid use and related mortalities in the United States between 2015 and 2019. The update to include questions on regular use of prescription opioids in 2015 [24] allowed the application of the extended method discussed in Section 3.4. We applied the model to compare the population-level progressions between non-Hispanic black and non-Hispanic white population.

5.1 | Data

We utilized multiple sources of data for the case study. We used the U.S. Mortality Multiple Cause Files [28] to extract opioid-induced mortality counts while we utilized the CDC Wonder interface [29] for general mortality rates and population counts. We derived other opioids use related counts and partial progression rates from the National Survey on Drug Use and Health

(NSDUH).[30] From each source, we collected separate data sets for non-Hispanic black and non-Hispanic white populations among those whose age was 12 or older recorded between 2015 and 2019.

We derived the opioid-induced mortality counts from the U.S. Mortality Multiple Cause Files.[28] The International Classification of Disease, Tenth Revision (ICD-10) codes reported in the files identify underlying causes of death and multiple causes of death. We followed the criteria used by Kumiko et al.[4] and used underlying cause-of-death codes X40-44 (unintentional), X60-64 (suicide), X85 (homicide), and Y10-Y14 (undetermined intent) to identify deaths caused by acute toxicity from drugs. Among the drug-related deaths, we identified those related to opioid using multiple cause-of-death codes: opium (T40.0), heroin (T40.1), natural opioid analgesics (T40.2), methadone (T40.3), synthetic opioid analgesics other than methadone including drugs such as fentanyl and tramadol (T40.4), or other and unspecified narcotics (T40.6). They inform the opioid-induced mortality counts, x_i for $i = 1, 2 \dots, 5$, to our proposed model. The general mortality rates, u_i for $i = 1, 2 \dots, 5$, were collected from the CDC Wonder.[29]

We relied on NSDUH[30] to derive the annual opioid misuse treatment initiate counts. We matched the recorded age groups of the survey participants and their responses to the question "How old were you when you first received treatment or counseling for your drug use?" among those who indicated that they had ever misused prescription opioids or use heroin. We used sample weights of the matched participants to estimate the yearly drug addiction treatment initiates. While we recognize the aggregation of individual age and the lack of specificity in the question may have introduced errors in the estimates, we assume they are negligible in in our aggregate model. We consider the estimates as the observed drug addiction treatment initiate counts, y_i for $i = 1, 2 \dots, 5$, in our model.

The NSDUH[30] also informed the annual counts for first-time illicit opioids users, w_i for $i = 1, 2, \dots, 5$. The survey asked a set of questions that indicated whether the respondent had misused either prescription opioids or used heroin for the first time in the past year.

Additionally, we used the NSDUH[30] to identify those who indicated they had ever used prescription opioids - both in accordance with a doctor's prescription and not, and those who had ever used heroin. We combined the information to inform $s_i^{(m)} \Big/ S_i^{(m)}$ for $m = -1, 2$ and $i = 1, 2 \dots, 5$ in the model.

5.2 | Priors

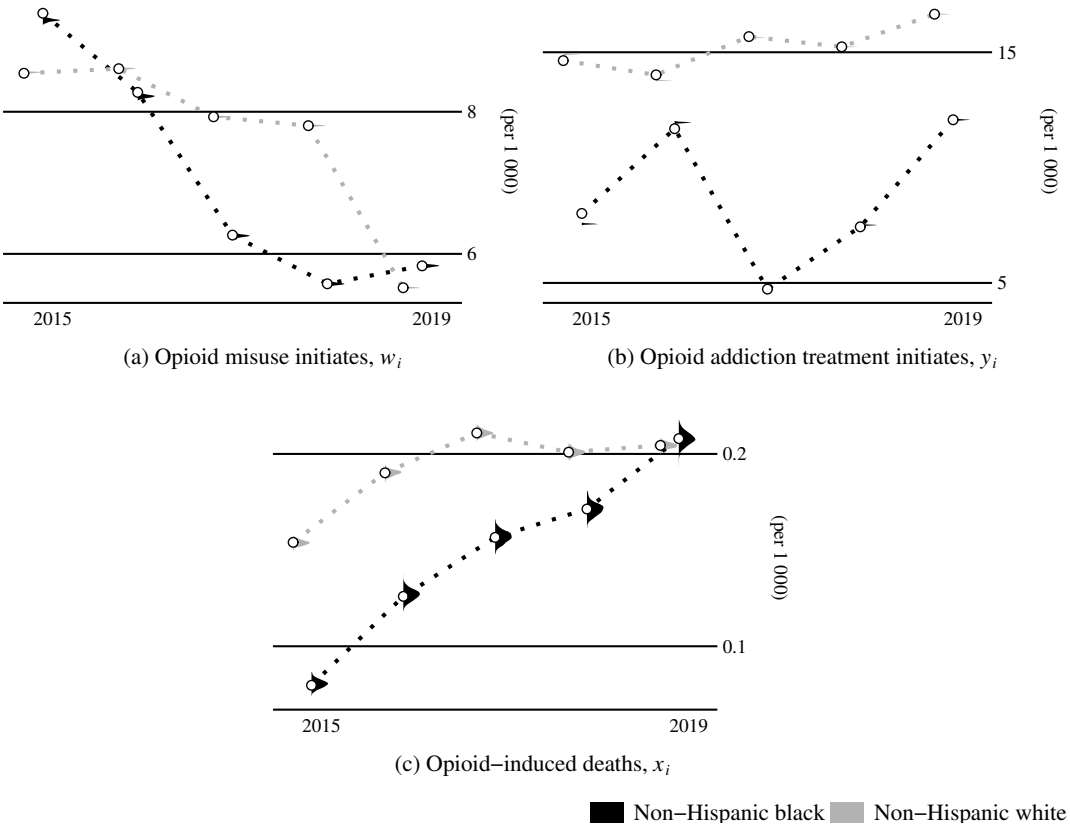
The proposed method requires a set of priors for a set as part of the Bayesian analysis. We specified weakly informative priors based on the available estimates and external information. We used a common set of priors for non-Hispanic black and non-Hispanic white populations to avoid introducing differences in the posterior distributions unseen in the data.

We first specified the initial expected incidences, $r_0^{(1)}$ and $r_0^{(2)}$. While there doesn't appear to be any publicly available data on first-time opioid prescriptions in the United States, Canadian Institute for Health Information[31] reported that 9.4% of the population in Ontario, Saskatchewan, and British Columbia in 2014. We assumed the proportion in the United States were similar and set $r_0^{(1)} = 0.1$. The NSDUH[30] estimated that 0.014% of the population in the United States aged 12 or older in 2015 first started using opioids illicitly. We set $r_0^{(2)} = 0.00015$. With ν_γ, we controlled the amount of deviations in $\gamma_0^{(1)} = \log(r_0^{(1)})$ and $\gamma_0^{(2)} = r_0^{(2)}$ over time. We set $\nu_\gamma = 1$ which allowed flexibility with probability of the annual expected incidences increasing by a factor of 2 or more to be approximately 0.157.

To specify the priors on the initial transition probabilities, we relied on the estimates from the NSDUH[30] and the observed opioid-induced mortality counts from the U.S. Mortality Multiple Cause Files[28] for the population in the United States aged 12 or older. Based on the estimates for 2015, we set $p_0^{(1)} = 0.015$, $p_0^{(2)} = 0.15$, and $p_0^{(3)} = 0.0005$. To control the amount of deviations, we set $\nu_\delta = 1$. The prior allowed the odd ratio of transition probabilities in comparison to the previous year to be larger than 2 with a probability of approximately 0.157.

We specified the priors for the initial prevalence of the first two states in our model based on the proportions of the population estimated from the NSDUH.[30] For the first state, we specified that the logit function of the proportion followed $N\left(0, 0.3^2\right)$. The estimated proportion from the survey in 2015 was approximately 0.5 and the prior allowed the proportion to be between 0.4 and 0.6 with a probability of approximately 0.826. Similarly, we specified the logit function of the proportion in the second state to be $N\left(-2, 1^2\right)$. The estimated proportion from the survey was approximately 0.1. The prior assumed a higher variance and allowed the proportion to be less than 0.3 with a probability of approximately 0.876.

Lastly, we assumed no direction in the biases introduced by the ratio estimates for the transition probabilities. We assumed normal with mean 0 and $\sigma^2 = 1$ truncated at -1 and 1 for both $\phi^{(1)}$ and $\phi^{(2)}$.



(a) Opioid misuse initiates, w_i

(b) Opioid addiction treatment initiates, y_i

(c) Opioid−induced deaths, x_i

Figure 8 Observed counts for non-Hispanic black and non-Hispanic white populations in the United States. Posterior predictive distributions based on 500 samples from the posterior samples for each group are shown to assess the model fits to the data.

5.3 | Results

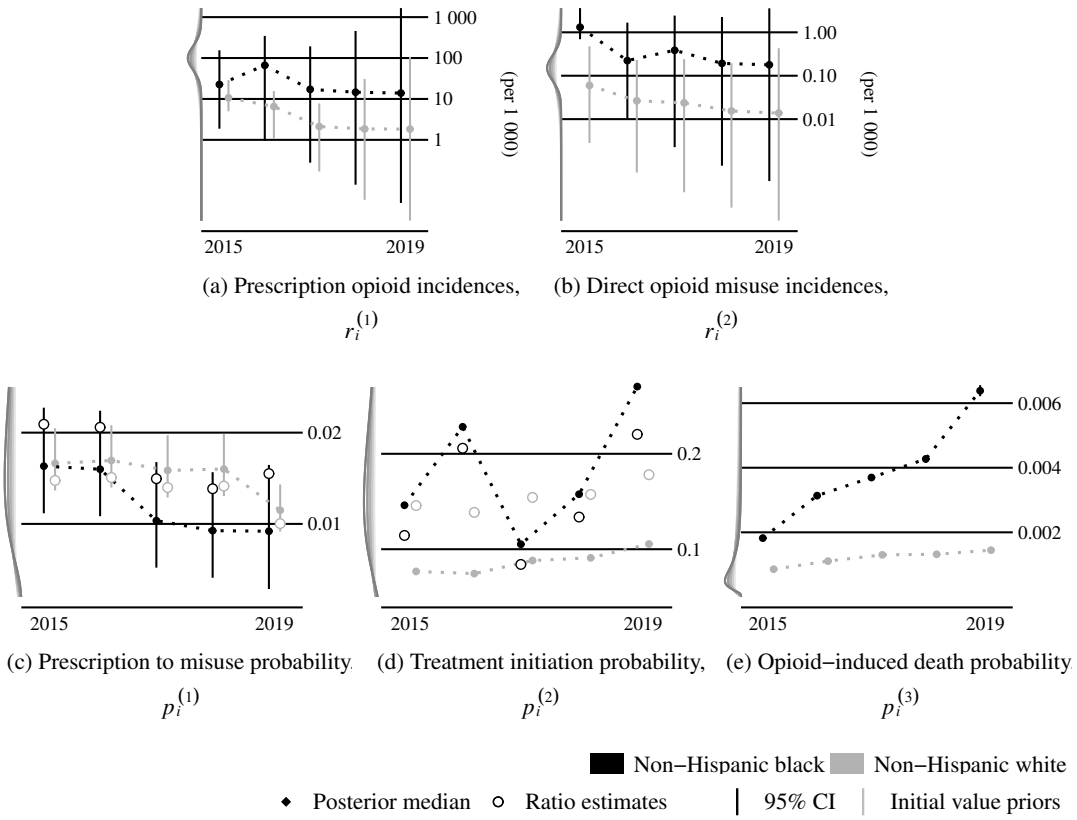
We fitted separate models for non-Hispanic black population and non-Hispanic white population using RStan's[25,26] implementation of No-U-Turn sampler.[27] Codes used for the fitting the models and generating the plots are available at https://github.com/mjmoon/bmm-opioid.

Figure 8 displays the count data used for fitting the model with the distribution of posterior predictive samples in black for the non-Hispanic black population and in gray for the non-Hispanic. To generate the posterior predictive distributions, we randomly selected 500 samples of their means from the posterior distributions, X_i^{post}, Y_i^{post}, and W_i^{post} and generated samples, x_i^{rep}, y_i^{rep}, and w_i^{rep}, based on the distributions described in Equation 10.

$$\begin{aligned} x_i^{\text{rep}} &\sim \text{Pois}\left(X_i^{\text{post}}\right) \\ y_i^{\text{rep}} &\sim \text{Pois}\left(Y_i^{\text{post}}\right) \\ w_i^{\text{rep}} &\sim \text{Pois}\left(W_i^{\text{post}}\right) \\ \text{for} \quad & i = 1, 2, \ldots, 5 \end{aligned} \tag{10}$$

The distributions of x_i^{rep} and w_i^{rep} align with the observed counts, x_i and w_i. The distributions y_i^{rep}, on the other hand, misses the observed counts for years between 2015 to 2017 although they track the trend over the years. By 2019, the predictive posterior distributions align for all three counts for both races. For all counts, the overall trends of the posterior predictive distributions align with the observed counts over time.

Figure 9 shows the posterior distributions of the incidences and the transition probabilities. The black markers represent non-Hispanic black population and the grey markers represent non-Hispanic white population. While the 95% credible intervals overlap, the incidence rates per 1 000 for opioid prescriptions and direct opioid misuse for non-Hispanic black population are consistently higher than those of the non-Hispanic white population. On the other hand, the model estimates higher transition



(a) Prescription opioid incidences, $r_i^{(1)}$

(b) Direct opioid misuse incidences, $r_i^{(2)}$

(c) Prescription to misuse probability $p_i^{(1)}$

(d) Treatment initiation probability, $p_i^{(2)}$

(e) Opioid−induced death probability $p_i^{(3)}$

Figure 9 Posterior medians with 95% credible intervals of the transition probabilities and incidences for non-Hispanic black and non-Hispanic white populations in the United States. The curves on the left margins represent the priors for the initial values based on 50th to 90th percentile values of σ_γ and σ_δ. The first two transition probabilities also include the ratio estimates used to inform the model.

probabilities from prescription use to misuse among non-Hispanic white population despite the ratio estimates from the NSDUH[30] showing the opposite. The model's estimates suggest that non-Hispanic black people were more likely to initiate opioids misuse directly from using illicit opioids. On the other hand, non-Hispanic white people initiate opioids misuse at a higher rate after taking prescription opioids.

Figure 9d suggest higher probabilities for non-Hispanic black population to initiate drug addiction treatments compared to non-Hispanic white population. Despite the higher treatment initiation probabilities, opioid-induced mortality rates among non-Hispanic black population increased more rapidly during the same time period as shown in Figure 8c. Figure 9e shows a similar trend in the probability of opioid-induced deaths among those who has not received treatments for non-Hispanic black population. The death probabilities are also estimated to be consistently higher among non-Hispanic black population.

DISCUSSION

We proposed a Bayesian back-calculation multistate framework that incorporates multiple sources of data to understand the progression to opioid-induced deaths. The framework allows reliable estimation of time-varying transition probabilities from opioid prescription to opioid misuse, from opioid misuse to addiction treatment, and from opioid misuse to opioid-induced deaths. It also estimates time-varying incidences of first time opioid prescription and opioid misuse. The Bayesian approach brings together data from multiple sources in a multistate model, and produces probabilistic estimates. Our proposed framework can depict the progression behaviour of opioid users at different stages while quantifying the uncertainties associated with the quantities estimated. While we were interested in the context of opioid use and related mortality, the framework would be broadly

applicable in cases where we are interested in the progression to observed outcomes with only partial information available for the intermediate steps.

The simulation studies presented in Section 4 demonstrate that the proposed framework is capable of retracting the unobserved incidences and interim transition probabilities with reasonable prior choices. They also reveal the potential challenges in practice. First, the posterior distributions of the incidences become less informative over a longer period of time with large variances as shown in Figure 5. Even within short periods, the variances may grow unexpectedly large. The case study on the opioid crisis in the United States exemplifies the issue as indicated by the 95% credible interval that span from less than 1 to 1 000 per 1 000 in Figure 9a. Using a different time-variant distribution model for the incidences such as higher order random walks may help address the issue. The simulation studies also reveal the lack of identifiability for the transition probability from opioid prescription to misuse also suffers as shown in Figure 7, which could be addressed by using informative priors.

The case study on the opioids crisis in the United States in Section 5 highlights the differences in how the crisis affected non-Hispanic black and non-Hispanic white populations. Specifically, the estimates suggest that the non-Hispanic black population faced higher probabilities of opioid-induced deaths conditional on using opioids in an illicit manner. The death probabilities were higher despite seeking treatments at higher rates. The difference in the type of initiating opioids between the two races may explain the disparity in the outcomes. Those who initiated opioids misuse with illicit opioids likely face higher risks of using illicitly manufactured synthetic opioids such as fentanyl. The trends also coincide with the increased number of deaths involving synthetic opioids while deaths involving prescription opioids remained stable during the same period.[32] Another possible contributing factor is the difference in quality of treatments. Non-Hispanic white opioid users likely receive higher quality of care and treatments for opioid abuse and dependence.[33] The difference in the type of initiating opioids and the disparity in care are likely causing more deaths among non-Hispanic black opioid users. Our framework was able to depict the differences that align with existing literature with the advantage of integrating them in a single, unified framework.

Our proposed framework depicts the aggregate behaviour at the population level. It could be extended to incorporate additional information. For example, the data sources used in Section 5[28,30] provide records by sex and age groups. Birrell et al.[23] demonstrates incorporating age as a covariate in a Bayesian back-calculation model to study HIV prevalence and incidences. A similar extension to our framework would allow age-specific estimates.

Another possible approach to extending the framework is extending the multistate model with more transitions and states. The 2015 update on the National Survey on Drug Use and Health[24] included addition of questions around prescription opioid use. This allowed us the model the prescription opioid users as a separate group in the multistate model shown in Figure 4. Continued improvements in monitoring of the opioids crisis would allow access to more information allowing the multistate model to be extended. For example, the literature suggests that a high proportion of those who receive treatments, including medicated treatments, on opioid addiction relapse to opioid misuse.[34,35,36] By incorporating additional information that, even partially, inform the reverse transition, the framework may be extended to provide estimates on the additional transition probability and improved estimates on the other transitions.

References

1. National Institute on Drug Abuse . Overdose death rates. 2022. U.S. Department of Health and Human Services.

2. Special Advisory Committee on the Epidemic of Opioid Overdoses . Opioid- and Stimulant-related Harms in Canada. 2022. Public Health Agency of Canada.

3. Alexander MJ, Kiang MV, Barbieri M. Trends in Black and White Opioid Mortality in the United States, 1979–2015. *Epidemiology* 2018; 29(5).

4. Lippold KM, Jones CM, Olsen EO, Giroir BP. Racial/Ethnic and Age Group Differences in Opioid and Synthetic Opioid-Involved Overdose Deaths Among Adults Ages ≥18 Years in Metropolitan Areas – United States, 2015-2017. *Morbidity and Mortality Weekly Report* 2019; 68(43).

5. Cochran BN, Flentje A, Heck NC, et al. Factors predicting development of opioid use disorders among individuals who receive an initial opioid prescription: Mathematical modeling using a database of commercially-insured individuals. *Drug and Alcohol Dependence* 2014; 138: 202–208. doi: 10.1016/j.drugalcdep.2014.02.701

6. Kiang MV, Basu S, Chen J, Alexander MJ. Assessment of Changes in the Geographical Distribution of Opioid-Related Mortality Across the United States by Opioid Type, 1999-2016.. *JAMA Netw Open* 2019; 2(2): e190040. doi: 10.1001/jamanetworkopen.2019.0040

7. Pitt AL, Humphreys K, Brandeau ML. Modeling Health Benefits and Harms of Public Policy Responses to the US Opioid Epidemic. *American Journal of Public Health* 2018; 108(10): 1394-1400. doi: 10.2105/AJPH.2018.304590

8. Battista NA, Pearcy LB, Strickland WC. Modeling the Prescription Opioid Epidemic. *Bulletin of Mathematical Biology* 2019; 81(7): 2258–2289. doi: 10.1007/s11538-019-00605-0

9. Brookmeyer R, Gail MH. A method for obtaining short-term prjections and lower boudns on the size of the AIDS epidemic. *Journal of the American Statistical Association* 1988; 83(402): 301-308.

10. Sweeting MJ, de Angelis D, Aalen OO. Bayesian back-calculation using a multi-state model with application to HIV. *Statistics ind Medicine* 2005; 24: 3991-4007.

11. Wu LT, Korper SP, Marsden ME, Lewis C, Bray RM. Use of Incidence and Prevalence in the Substance Use Literature: A Review. tech. rep., Office of Applied Studies, Substance Abuse and Mental Health Services Admnistration, U.S. Department of Health & Human Services; Rockville, MD: 2003.

12. Alarcón dR. The spread of heroin abuse in a community. *Bulletin on Narcotics* 1969; 21(3): 17-22.

13. Mackintosh DR, Stewart GT. A mathematical model of a heroin epidemic: implications for control policies. *Journal of Epidemiology and Community Health* 1979; 33: 299-304.

14. Rossi C. The role of dynamic modellingin drug abuse epidemiology. *Bulletin on Narcotics* 2002; 54(1): 33-44.

15. White E, Comiskey C. Heroin epidemics, treatment and ODE modelling. *Mathematical Biosciences* 2007; 208: 312-324.

16. Hethcote HW. The Mathematics of Infectious Diseases. *SIAM Review* 2000; 42(4): 599-653.

17. Rossi C. Operational models for epidemics of problematic drug use: the Mover-Stayer approach to heterogeneity. *Socio-Economic Planning Sciences* 2004; 38(1): 73-90.

18. Djilali S, Touaoula TM, Miri SEH. A heroin epidemic model: Very general non linear incidence, treat-age, and global stability. *Acta Applicandae Mathematicae* 2017; 152(1): 172-194.

19. de Angelis D, Hickman M, Yang S. Estimatingg long-term trends in the incidence and prevalence of opiate use/injecting drug use and the number of former users: Back-calculation methods and opiate overdose deaths. *American Journal of Epidemiology* 2004; 160(10): 994-1004.

20. Sánchez-Niubò A, Aalen OO, Domingo-Salvany A, Amundsen EJ, Fortiana J, Røysland K. A multi-state model to estimate incidence of heroin use. *BMC Medical Research Methodology* 2013; 13(1): 4. doi: 10.1186/1471-2288-13-4

21. Aalen OO, Farewell VT, de Angelis D, Day NE, Gill ON. A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: Application to AIDS prediction in England and Wales. *Statistics ind Medicine* 1997; 16: 2191-2210.

22. Birrell PJ, Chadborn TR, Gill ON, Delpech VC, de Angelis D. Estimating trends in incedence, time-to-diagnosis and undiagnosed prevalence using a CD4-based Bayesian back-calculation. *Statistical Communications in Infectiouus Diseases* 2012; 4(1).

23. Brizzi F, Birrell PJ, Plummer MT, et al. Extending Bayesian back-calculation to estimate age and time specific HIV incidence. *Lifetime Dat Analysis* 2019; 25: 757-780.

24. Substance Abuse and Mental Health Services Administration . 2015 National Survey on Drug Use and Health - Summary of the Effects of the 2015 NSDUH Questionnaire Redesign: Implications for Data Users. 2016.

25. Stan Development Team . RStan: the R interface to Stan. 2023. R package version 2.21.8.

Table 1 True parameter values for generating the simulation data.

Parameter	Value
$r_0^{\text{TRUE,(m)}}$	$\begin{cases} 10\,000 & m = 1 \\ 1\,000 & m = 2 \end{cases}$
$\sigma_\gamma^{\text{TRUE}}$	0.5
$p_0^{\text{TRUE,(m)}}$	$\begin{cases} 0.05 & m = 1 \\ 0.1 & m = 2 \\ 0.005 & m = 3 \end{cases}$
$\sigma_\delta^{\text{TRUE}}$	0.05
u^{TRUE}	0.02

26. Stan Development Team . Stan modeling language users guide and reference manual. 2022. Version 2.21.0.

27. Homan MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 2014; 15(1): 1593–1623.

28. Centers for Disease Control and Prevention, National Center for Health Statistics . Mortality Multiple Cause Files. 2022. https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Mortality_Multiple.

29. Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System . Mortality 1999-2020 on CDC WONDER Online Database. 2021. Data are from the Multiple Cause of Death Files, 1999-2020, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed at http://wonder.cdc.gov/mcd-icd10.html on Mar 16, 2023 2:29:51 AM.

30. Substance Abuse and Mental Health Services Administration . National Survey on Drug Use and Health (NSDUH). 2023.

31. Canadian Institute for Health Information . Opioid Prescribing in Canada: How Are Practices Changing?. 2019.

32. Wilson N, Kariisa M, Seth P, IV HS, Davis NL. Drug and Opioid-Involved Overdose Deaths — United States, 2017–2018. *Morbidity and Mortality Weekly Report* 2020; 69(11): 290–297.

33. Santoro TN, Santoro JD. Racial Bias in the US Opioid Epidemic: A Review of the History of Systemic Bias and Implications for Care. 2018; 10(12): e3733. doi: 10.7759/cureus.3733

34. Bailey GL, Herman DS, Stein MD. Perceived Relapse Risk and Desire for Medication Assisted Treatment among Persons Seeking Inpatient Opiate Detoxification. *Journal of Substance Abuse Treatment* 2013; 45(3): 302–305. doi: 10.1016/j.jsat.2013.04.002

35. Weiss RD, Rao V. The Prescription Opioid Addiction Treatment Study: What have we learned. *Drug and Alcohol Dependence* 2017; 173: S48-S54. Prescription Opioids: new perspectives and research on their role in chronic pain management and addictiondoi: https://doi.org/10.1016/j.drugalcdep.2016.12.001

36. Smyth B, Barry J, Keenan EE, Ducray K. Lapse and relapse following inpatient treatment of opiate dependence. *The Irish Medical Journal* 2016; 109(1).

How to cite this article: Moon M.J. and M. Alexander (2023), Modelling transitions of opioid usage, addiction, and fatal overdoses using a Bayesian multistate model, *Statistics in Medicine*, .