Psychological Methods

Bayesian Continuous-Time Hidden Markov Models With Covariate Selection for Intensive Longitudinal Data With Measurement Error

Mingrui Liang, Matthew D. Koslovsky, Emily T. Hébert, Darla E. Kendzor, Michael S. Businelle, and Marina Vannucci

Online First Publication, December 20, 2021. http://dx.doi.org/10.1037/met0000433

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Methods

ISSN: 1082-989X

https://doi.org/10.1037/met0000433

Bayesian Continuous-Time Hidden Markov Models With Covariate Selection for Intensive Longitudinal Data With Measurement Error

Mingrui Liang[1], Matthew D. Koslovsky[2], Emily T. Hébert[3], Darla E. Kendzor[4], Michael S. Businelle[4], and Marina Vannucci[1]

[1] Department of Statistics, Rice University

[2] Department of Statistics, Colorado State University

[3] Department of Health Promotion and Behavioral Sciences, University of Texas Health Science Center at Austin (UTHealth) School of Public Health

[4] Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center

Abstract

Intensive longitudinal data collected with ecological momentary assessment methods capture information on participants' behaviors, feelings, and environment in near real-time. While these methods can reduce recall biases typically present in survey data, they may still suffer from other biases commonly found in self-reported data (e.g., measurement error and social desirability bias). To accommodate potential biases, we develop a Bayesian hidden Markov model to simultaneously identify risk factors for subjects transitioning between discrete latent states as well as risk factors potentially associated with them misreporting their true behaviors. We use simulated data to demonstrate how ignoring potential measurement error can negatively affect variable selection performance and estimation accuracy. We apply our proposed model to smartphone-based ecological momentary assessment data collected within a randomized controlled trial that evaluated the impact of incentivizing abstinence from cigarette smoking among socioeconomically disadvantaged adults.

Translational Abstract

Continuous-time hidden Markov models enable researchers to study the relations between risk factors and outcomes repeatedly measured at unbalanced, unequally spaced assessment times while accommodating measurement error. Despite extensive implementation of variable selection methods designed to identify potential relations in exploratory settings for hypothesis generation, currently none have been applied to this class of models commonly found in psychological research. To fill this gap, we develop a Bayesian continuous-time hidden Markov model with variable selection priors and provide a flexible R package to facilitate the application of our method in practice. We showcase the variable selection performance and estimation accuracy of our method on simulated data and apply it to intensive longitudinal data collected in a smoking cessation trial to identify potential risk factors associated with smoking behaviors after a quit attempt.

Keywords: Bayesian variable selection, hidden Markov model, measurement error, mHealth, smoking cessation

Supplemental materials: https://doi.org/10.1037/met0000433.supp

Psychological researchers often use mobile health (mHealth) methods to monitor participants in their natural environments to improve health-related outcomes through behavioral change. For example, clinical psychologists have used smartphones to collect intensive longitudinal data via ecological momentary assessments (EMAs), which aim to capture psychological, emotional, and environmental factors that may relate to a behavioral outcome in near real-time. The high-temporal resolution of these data helps reduce

Mingrui Liang https://orcid.org/0000-0002-9507-9713
Matthew D. Koslovsky https://orcid.org/0000-0001-5144-2042
Emily T. Hébert https://orcid.org/0000-0001-5922-164X
Michael S. Businelle https://orcid.org/0000-0002-9038-2238
Marina Vannucci https://orcid.org/0000-0002-7360-5321

Mingrui Liang and Matthew D. Koslovsky shared equal contribution.

This research was primarily supported by National Cancer Institute (NCI) Grant R01CA197314 awarded to Darla E. Kendzor. Additional support was provided by the Oklahoma Tobacco Settlement Endowment Trust (TSET; Grant R21-02) and NCI Support Grant P30CA225520 awarded to the Stephenson Cancer Center. Results from this article have not been previously disseminated; however, data from the parent study have been presented at the Society for Research on Nicotine and Tobacco and the Society of Behavioral Medicine Annual Meetings. Emily T. Hébert was supported by National Institute on Drug Abuse Grant R00DA046564.

Correspondence concerning this article should be addressed to Matthew D. Koslovsky, Department of Statistics, Colorado State University, 207 Statistics Building, 851 Oval Drive, Fort Collins, CO 80523, United States. Email: matt.koslovsky@colostate.edu

memory errors and recall biases typically present in self-reported data (Shiffman et al., 2008). More recently, researchers have used smartphones and other wearable devices to passively capture behavioral and health-related information, including patterns of activity and movement via geospatial measurements and accelerometers as well as physiological factors, such as heart rate, skin conductance, and voice characteristics (Bentley et al., 2019). As a result, researchers are able to continuously collect multimodal, intensive longitudinal data on subjects' potential risk factors and outcomes from a third-person observational perspective which enables ecological validity (Nelson & Allen, 2018). These methods have not only helped researchers better understand complex psychological and behavioral processes, but additionally have enhanced their ability to design, evaluate, and deliver tailored intervention strategies based on a subject's risk profile at critical moments throughout the assessment period (Businelle et al., 2016; Hébert et al., 2020; Heron & Smyth, 2010; Nahum-Shani et al., 2018; Rehg et al., 2017).

Despite the potential utility of mHealth methods in psychological research, remote data collection presents a challenge for the verification of study outcomes. For example, within the context of smoking cessation research, accurate reporting of smoking status is important not only for verifying smoking cessation intervention outcomes, but for identifying moments of high risk that could serve as targets for just-in-time adaptive interventions (Nahum-Shani et al., 2018). However, smokers may underreport smoking behaviors due to social desirability bias (Patrick et al., 1994). Thus, recent research has explored methods to biochemically validate smoking status remotely, including direct measurement (e.g., using personal carbon monoxide monitors; Kendzor et al., 2020; McClure et al., 2018) and indirect measurement via sensors (e.g., heart rate; Herbec et al., 2020). However, sensor data are often messy, typically require extensive preprocessing, which can smooth out signals, and have accuracy levels that depend on the type of metric being measured (Bentley et al., 2019; Schukat et al., 2016). Even with high-quality data, it is difficult to determine or map which factors are attributed to a change in a particular physiological measurement (Schukat et al., 2016). By ignoring potential measurement error, researchers may obtain biased estimates for the relations between outcomes and risk factors when fitting regression models (Carroll et al., 2006; Yen & Chen, 2018).

One strategy for accommodating potential measurement error and improving the internal validity of the study is through data integration or fusion (Kumar et al., 2013; Mitchell, 2007; Rehg et al., 2017). With this approach, researchers may couple actively collected data with passively collected data to provide objective information that contextualizes and validates self-reported momentary behavior and mood (Bertz et al., 2018; Bond et al., 2014; Kumar et al., 2013). In theory, reliable and accurate sensor data could even be used to replace self-reported data, reducing patient burden while increasing the temporal resolution of the data (Bertz et al., 2018; Blaauw et al., 2016).

More commonly, researchers accommodate potential measurement error through their analytical approach via latent variable models (Bandalos, 2018). A variety of statistical methods have been developed to accommodate measurement error for both continuous and discrete outcomes (Carroll et al., 2006; Gustafson, 2003; Li & Vuong, 1998; Yen & Chen, 2018). One of the most popular approaches for handling measurement error in longitudinal psychological studies is latent or hidden Markov models (HMMs; Yen & Chen, 2018). These methods are an extension of the class of multistate Markov (MSM) models, which were designed to model subjects as they transition between discrete states over time (Kalbfleisch & Lawless, 1985; Ma et al., 2015, 2018; Peng et al., 2019). HMMs extend MSM models by assuming that the repeated data collected on subjects are realizations (emissions) of an unobserved process, characterized by discrete latent states, which subjects transition through over time. As such, these methods are commonly applied to longitudinal data collected in psychological studies to investigate constructs that are difficult to measure directly and to accommodate potential measurement errors.

A primary objective in longitudinal data analyses is identifying or reaffirming complex relations between risk factors associated with outcomes over time (Walls, 2013). While there are numerous MSM and HMM tutorials and software packages available that provide guidelines for performing model selection (Jackson, 2011; Kaplan, 2008; Visser, 2011; Visser et al., 2002), little methodological work has been devoted to developing automated variable selection methods for hypothesis generation. Recently, Reulen and Kneib (2016) and Sennhenn-Reulen and Kneib (2016) designed variable selection methods for MSM models when the exact time of transition is known. Additionally, Koslovsky et al. (2018) developed a variable selection method for MSM models using expectation-maximization for interval-censored data (Dempster et al., 1977). While these methods are suitable for the longitudinal data sets typically found in psychological research (i.e., unbalanced, randomly spaced assessment times), they ignore potential measurement error. Recently, researchers have designed variable selection methods for discrete-time hidden Markov models using Bayesian and frequentist techniques (Kang et al., 2019; Rashid et al., 2014; Spezia, 2020). However, these methods are inappropriate for unequally spaced assessment times and are therefore not applicable to a majority of psychological studies.

The primary objective of this article is to develop a novel, fully Bayesian variable selection approach for *continuous-time* hidden Markov models. The proposed approach accommodates potential biases by simultaneously identifying risk factors for subjects transitioning between discrete latent states as well as risk factors associated with them potentially misreporting their true behavior. The development of the proposed model is motivated by intensive longitudinal data collected in the PREVAIL II study, a randomized controlled trial designed to evaluate the efficacy of offering abstinence-contingent financial incentives for smoking cessation among socioeconomically disadvantaged adults. In this study, participants were repeatedly prompted with ecological momentary assessments of their current environment, affect, behaviors, social interactions, and recent smoking behaviors on a study-provided smartphone. Assessments were prompted five times per day over a four week period. Smoking abstinence, defined as expired carbon monoxide less than or equal to 6 ppm (based on current recommendations; Benowitz et al., 2020), was biologically verified at in-person visits every week. Instead of modeling the observed transitions between momentary smoking states directly, we assume a hidden layer for the true response sequence and consider the reported behavior as an emission process from the truth to account for potential reporting biases and measurement errors. By performing variable selection in both levels of the model, we are able to learn which risk factors are associated with transitions between discrete latent smoking states, and also which risk factors are associated with subjects misreporting their smoking behaviors. We use simulated data to demonstrate the variable selection performance and estimation

accuracy of our method in various settings, including how ignoring potential measurement error can negatively affect results. While developed for intensive longitudinal data collected in mHealth studies for smoking cessation, the proposed method can be used to perform variable selection in research settings where a two-state continuous-time hidden Markov model with binary emissions is suitable. To help researchers use our proposed approach in practice, we provide R code and an accompanying vignette applying our method to simulated data.

In the Methods section, we introduce the proposed Bayesian variable selection method for both continuous-time multistate and hidden Markov models. In the Case Study, we apply our model to intensive longitudinal data collected via smartphone to investigate the benefits of providing smoking incentives to promote smoking cessation. In the Simulation Study, we evaluate the selection and estimation performance of our proposed model on simulated data in various settings and compare to the MSM model. We conclude with final remarks.

Method

To introduce HMMs, we first describe how to construct transition probabilities for a continuous-time multistate Markov model, which could be used to study the relations between risk factors and transitions ignoring potential measurement error. We then extend this model to account for measurement error via the HMM and describe how to incorporate variable selection through prior specification using a Bayesian framework. In this work, we focus on two-state models based on our application to smoking behaviors.

Modeling Hidden and Observed States

Let $y_{ij} \in \{0, 1\}$ represent the observed binary outcome of the i^{th} subject, $i = 1, 2, \ldots, \dot{N}$, at the j^{th} assessment time, $j = 1, \ldots, n_i$. For example, y_{ij} may capture a subject's reported smoking behavior (smoking or nonsmoking), activity status (active or sedentary), or medication adherence (yes or no). This formulation allows for unbalanced and unequally spaced assessment times across subjects. We define $h_{ij} \in \{0, 1\}$ as the true, latent or hidden state for individual i at the j^{th} assessment time.

Transition Probabilities

Transition probabilities are modeled using a continuous-time multistate Markov model framework, similar to Kalbfleisch and Lawless (1985) and Koslovsky et al. (2018). In this two-state setting, λ and μ represent the positive transition rates from $0 \rightarrow 1$ and $1 \rightarrow 0$, respectively (Cox & Miller, 2017). The transition rate or intensity matrix $\boldsymbol{Q}$ for a homogenous Markov process can be represented as

$$\begin{array}{ccc} & \textit{current} & \textit{state} \\ & 0 & 1 \\ \textit{previous} \quad 0 & -\lambda & \lambda \\ \textit{state} \quad 1 & \mu & -\mu \end{array}$$

where the elements represent the instantaneous rate at which the continuous-time Markov chain transitions between states. The corresponding transition probability matrix $\boldsymbol{P} = \exp(\boldsymbol{Q}\delta_{ij})$ is defined as

$$\begin{array}{ccc} & \textit{current} & \textit{state} \\ & 0 & 1 \\ \textit{previous} \quad 0 & P(0,0 \mid \delta_{ij}) & P(0,1 \mid \delta_{ij}) \\ \textit{state} \quad 1 & P(1,0 \mid \delta_{ij}) & P(1,1 \mid \delta_{ij}) \end{array}$$

where $P(0,1|\delta_{ij})$ is the probability of transitioning from state 0 to state 1 given $\delta_{ij} = t_{i,j} - t_{i,j-1}$, the change in time between the current and previous assessment. Under this assumption, the transition probabilities have closed-from solutions, following Pinsky and Karlin (2010),

$$P(0,1 \mid \delta_{ij}) = 1 - P(0,0 \mid \delta_{ij}) = \frac{\lambda}{\lambda + \mu}[1 - \exp(-(\lambda + \mu)\delta_{ij})]$$

and

$$P(1,0 \mid \delta_{ij}) = 1 - P(1,1 \mid \delta_{ij}) = \frac{\mu}{\lambda + \mu}[1 - \exp(-(\lambda + \mu)\delta_{ij})].$$

Note that these transition probabilities can be used to construct the likelihood function for the observed outcomes $\boldsymbol{y}$ in an MSM model, as well as the probability of transitioning between hidden states $\boldsymbol{h}$ in a continuous-time HMM. In either case, the model is equipped to handle unbalanced, unequally spaced assessment times in which the exact time of transition is unknown (i.e., transition times are interval-censored).

In order to identify covariates associated with transitions between discrete states, we embed covariate dependent transition rates, similar to Jones et al. (2006) and Koslovsky et al. (2018). Let $\boldsymbol{x}_{ij} = (x_{ij1}, \ldots, x_{ijp})'$ represent a p-dimensional vector of observed covariates or risk factors, for the i^{th} subject at the j^{th} assessment time. Then, the transition rates λ and μ can be redefined as $\lambda = \lambda_{ij} = \exp(\lambda_0 + \dot{\boldsymbol{x}}_{ij}'\boldsymbol{\beta}_\lambda)$ and $\mu = \mu_{ij} = \exp(\mu_0 + \dot{\boldsymbol{x}}_{ij}'\boldsymbol{\beta}_\mu)$, where λ_0 and μ_0 are interpreted as the log baseline hazards for transitioning between states, and $\dot{\boldsymbol{x}}_{ij} = (\dot{x}_{ij1}, \ldots, \dot{x}_{ij\dot{p}})'$ is a subset of $\boldsymbol{x}_{ij}$ with $\dot{p} \leq p$. Similarly, $\boldsymbol{\beta}_\lambda = (\beta_{\lambda,1}, \ldots, \beta_{\lambda,\dot{p}})'$ and $\boldsymbol{\beta}_\mu = (\beta_{\mu,1}, \ldots, \beta_{\mu,\dot{p}})'$ represent log hazard ratios for their corresponding covariates. Under this construction, time-varying covariates are assumed constant between assessment times, and the model can readily handle time-invariant or baseline covariates as well. Additionally, each covariate is allowed to have different relations with each transition, and the estimated effects are fixed across time. Additionally, the previous value of time-varying covariates could be used to assess transitions, however, longer lags in covariate patterns would violate the Markovian assumption. Lastly, initial state probabilities may be treated as nuisance parameters when inference is focused on transitions, similar to our setting (Benoit et al., 2016), assumed to be subject-specific, or specified to depend on baseline covariates (Zhou et al., 2020). We assume similar initial state probabilities, π_1, across all subjects,

$$\pi_1(h_{i1} = 1) = 1 - \pi_1(h_{i1} = 0) = \pi.$$

In our simulation study below, we demonstrate the model's insensitivity to this assumption on variable selection performance for transitions and emissions.

Emission Probabilities

We model the relation between observed (or reported) states (typically referred to as emission data in the context of HMMs) and hidden states using a logistic regression framework. We assume that each observed outcome, y_{ij}, is independent of all previous observations and the hidden processes prior to time t_{ij}, conditional on the hidden state, h_{ij}. Within the logistic framework, we model "success" as a truthful, or accurate, response, similar to Bureau et al. (2003),

$$f(y_{ij} \mid h_{ij} = 0) = \omega_{0,ij}^{1-y_{ij}}(1 - \omega_{0,ij})^{y_{ij}}$$

$$f(y_{ij} \mid h_{ij} = 1) = \omega_{1,ij}^{y_{ij}}(1 - \omega_{1,ij})^{1-y_{ij}},$$

where $\omega_{0,ij}$ and $\omega_{1,ij}$ represent the probability of a matched hidden and observed state (i.e., a measurement collected without error). While this framework accounts for potential measurement error in the reported outcomes, it is important to note that it does not differentiate between different sources of measurement error. To account for covariate effects on the emission process, the emission probabilities can be modeled using a logit function as follows,

$$logit(\omega_{\nu,ij}) = \log\left(\frac{\omega_{\nu,ij}}{1 - \omega_{\nu,ij}}\right) = \beta_{\nu,0} + \ddot{\boldsymbol{x}}_{ij}'\boldsymbol{\beta}_{\nu}, \quad \nu = 0, 1,$$

where $\ddot{\boldsymbol{x}}_{ij} = (\ddot{x}_{ij1}, \ldots, \ddot{x}_{ij\ddot{p}})'$ with $\ddot{p} \leq p$. Note that with this parameterization, our model allows different sets of covariates to be associated with transitions and emissions. Here, $\beta_{\nu,0}$ and $\boldsymbol{\beta}_{\nu} = (\beta_{\nu,1}, \ldots, \beta_{\nu,\ddot{p}})'$ are interpreted as the baseline and covariate specific log odds ratios for properly reporting the outcome, respectively. A graphical representation of the relationship between hidden and observed states is presented in Figure 1.

Prior Specification

To identify risk factors associated with transitions between discrete states and accurately reporting outcomes, we impose spike-and-slab priors for each of the K regression coefficients $\boldsymbol{\beta}' = (\boldsymbol{\beta}_{\lambda}', \boldsymbol{\beta}_{\mu}', \boldsymbol{\beta}_{0}', \boldsymbol{\beta}_{1}')$, where

$$\beta_k \sim \gamma_k N(0, v_k) + (1 - \gamma_k)\delta_0(\beta_k), \quad k = 1, \ldots, 2\dot{p} + 2\ddot{p} = K$$

(Brown et al., 1998; George & McCulloch, 1997). In general, spike-and-slab prior distributions are composed of a mixture of a diffuse distribution (slab) and a Dirac delta function at zero (spike), δ_0. While other sparsity inducing priors are available (Van Erp et al., 2019), we chose the spike-and-slab prior since it forces nonactive terms to zero and explicitly provides inference on each covariate's posterior probability of inclusion. Here, we assume the slab follows a $N(0,v_k)$, where v_k is diffuse to allow active covariates to be freely estimated. A latent inclusion indicator, $\gamma_k \in \{0,1\}$, is assigned to each regression coefficient, which represents whether or not the corresponding covariate is excluded or included in the model. We assume that each of the γ_k in the $(2\dot{p} + 2\ddot{p})$ -dimensional inclusion indicator vector $\boldsymbol{\gamma}' = (\boldsymbol{\gamma}_{\lambda}', \boldsymbol{\gamma}_{\mu}', \boldsymbol{\gamma}_{0}', \boldsymbol{\gamma}_{1}')$ follows a Bernoulli distribution with sparsity parameter $\theta_k \in [0,1]$,

$$f(\gamma_k \mid \theta_k) = \theta_k^{\gamma_k}(1 - \theta_k)^{1-\gamma_k}.$$

By assuming a conjugate Beta(a_k, b_k) distribution for the θ_k, we can integrate out θ_k such that the inclusion indicator's marginal prior distribution is

$$\begin{aligned}\pi(\gamma_k) &= \frac{\text{Beta}(\gamma_k + a_k, 1 - \gamma_k + b_k)}{\text{Beta}(a_k, b_k)} \\ &= \frac{\Gamma(\gamma_k + a_k)\Gamma(1 - \gamma_k + b_k)\Gamma(a_k + b_k)}{\Gamma(1 + a_k + b_k)\Gamma(a_k)\Gamma(b_k)}.\end{aligned}$$

The hyperparameters a_k and b_k control the sparsity of the model. The beta-binomial prior is commonly used to model inclusion indicators when no other information regarding a covariate's inclusion in the model is known. In settings where additional covariate information or functional relationships between covariates are present, logistic and Markov random field priors are often employed (Stingo et al., 2010; Stingo & Vannucci, 2011). Note that some or all of the covariates can be forced into both levels of the model by fixing their corresponding inclusion indicator $\gamma_k = 1$. To finish the prior specification of our model, we let the baseline transition and emission variables λ_0, μ_0, $\beta_{0,0}$, and $\beta_{1,0}$ follow a normal distribution with mean 0 and variance v_0, where v_0 is diffuse so that the intercept terms are freely estimated.

Posterior Inference

For posterior inference, we implement a Metropolis-Hastings within Gibbs algorithm. The full joint model is defined as

$$\begin{aligned}&p(\Theta \mid \boldsymbol{y}, \boldsymbol{h}, \boldsymbol{X}) \\ &\propto \prod_{i=1}^{\dot{N}} p(\boldsymbol{h}_i, \boldsymbol{y}_i \mid \boldsymbol{x}_i, \Theta)\pi(\boldsymbol{\beta} \mid \boldsymbol{\gamma})\pi(\boldsymbol{\gamma})\pi(\lambda_0)\pi(\mu_0)\pi(\beta_{0,0})\pi(\beta_{1,0}),\end{aligned}$$

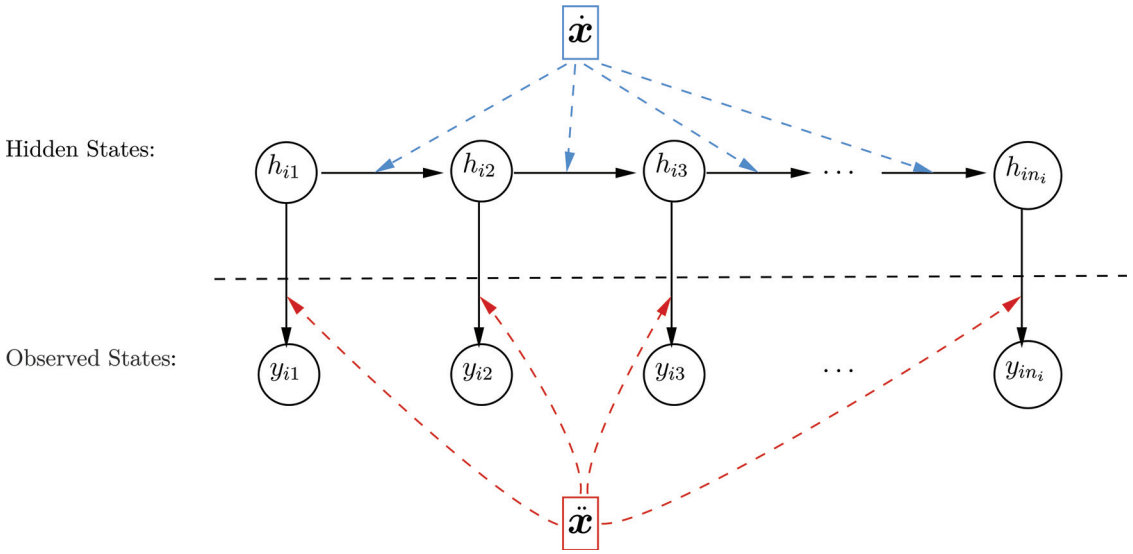
where $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \lambda_0, \mu_0, \beta_{0,0}, \beta_{1,0}\}$ and $\boldsymbol{X}$ represents the $\sum_{i=1}^{\dot{N}} n_i \times p$ matrix of covariates, with $\boldsymbol{x}_i$ containing the i^{th} subject's observed covariates. The joint likelihood of the observed data, $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{ij}, \ldots, y_{in_i})'$, and hidden state sequence, $\boldsymbol{h}_i = (h_{i1}, \ldots, h_{ij}, \ldots, h_{in_i})'$, is

$$p(\boldsymbol{h}_i, \boldsymbol{y}_i \mid \boldsymbol{x}_i, \Theta) = \pi_1(h_{i1})f(y_{i1} \mid h_{i1}, \ddot{\boldsymbol{x}}_{i1})\prod_{j=2}^{n_i} q(h_{ij-1}, h_{ij} \mid \delta_{ij}, \dot{\boldsymbol{x}}_{ij})f(y_{ij} \mid h_{ij}, \ddot{\boldsymbol{x}}_{ij}).$$

We use the Pólya-Gamma augmentation technique of Polson et al. (2013) to efficiently sample the posterior distribution for emission parameters in the logistic regression model because it maintains interpretability of regression coefficients as log odds ratios. Hidden states are sampled using the scaled forward-backward algorithm (Scott, 2002). Inclusion indicators and regression coefficients are jointly updated using an Add/Delete step, following Savitsky et al. (2011).

The MCMC sampler used to implement the proposed HMM, as well as an MSM model, is outlined below in Algorithm 1. A more detailed description of the MCMC steps as well as model derivations are provided in the online supplemental materials. After burn-in and thinning, the remaining samples obtained from running Algorithm 1 for M iterations are used for inference. To

Figure 1
Graphical Representation of the Assumed Dependence Structure in the Hidden Markov Model for the Observed (y_{ij}) and Hidden (h_{ij}) States for Subject i



Note. See the online article for the color version of this figure.

determine a risk factor's inclusion in the model, its marginal posterior probability of inclusion (MPPI) is determined by calculating the average of its respective inclusion indicator's MCMC samples. Commonly, covariates are included in the model if their MPPI exceeds .50, often referred to as the median model approach (Barbieri & Berger, 2004), or a Bayesian false discovery rate (BFDR) threshold, which controls for multiplicity (Newton et al., 2004).

Algorithm 1: MCMC Algorithm Used for MSM and HMMs

1. Input data $\boldsymbol{y}$ and $\boldsymbol{X}$
2. Initiate $\boldsymbol{\beta}$, γ, π, μ_0, λ_0, $\beta_{0,0}$, and $\beta_{1,0}$
3. Set *HMM* equal to True or False for hidden or multistate Markov model, respectively.
4. **for** iteration $m = 1, \ldots, M$ **do**
5. **if** *HMM* **then**
6. Run a forward-backward algorithm to simulate $\boldsymbol{h}$ given current $\boldsymbol{\beta}$ and γ (Scott, 2002)
7. **end if**
8. Jointly update $\boldsymbol{\beta}_\lambda$, $\boldsymbol{\beta}_\mu$, γ_λ, and γ_μ with a Between step (Savitsky et al., 2011)
9. Perform a Within Step for $\boldsymbol{\beta}_\lambda$, $\boldsymbol{\beta}_\mu$, μ_0, and λ_0 with a Metropolis-Hastings update (Savitsky et al., 2011)
10. **if** *HMM* **then**
11. Jointly update $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, γ_0, and γ_1 with a Between step (Savitsky et al., 2011)
12. Perform a Within Step for $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, $\beta_{0,0}$, and $\beta_{1,0}$ with a Pólya-Gamma update (Polson et al., 2013)
13. **end if**
14. **end for**

Label Switching

HMMs are subject to label switching, as the likelihood of the data is invariant to label permutations (Redner & Walker, 1984; Scott, 2002). Label switching occurs due to the nonidentifiability of the hidden components, meaning the label assigned to each cluster (true smoking or nonsmoking status in our case study) is meaningless and can cause inferential challenges when labels for distinct clusters swap throughout the MCMC iterations. In practice, label switching is typically diagnosed via jumps in trace plots in addition to multimodal density plots for parameters. When the proportion of observations that are misreported is known to be low, another way to check for this phenomenon in two-state models is to assess the correspondence between estimated and observed states. There are a variety of techniques to handle label switching, including parameter constraints to break the symmetry of the likelihood and relabeling algorithms (Jasra et al., 2005; Papastamoulis, 2015; Stephens, 2000). In simulation, we only observed label switching when our method was applied to data with a high proportion of misreported outcomes (i.e., low baseline emission probabilities). While a somewhat crude approach, we found that simply seeding the MCMC algorithm at the maximum likelihood estimate solved any label switching issues in low emission probability settings. In our contributed R package, users are able to accommodate label switching in their analyses by constraining the model space so that one of intercept terms in the transition (i.e., $\lambda 0$ and μ_0) and/or emission (i.e., $\beta_{0,0}$ and $\beta_{1,0}$) levels of the model is higher than the other, in addition to controlling where the model is initialized. For example in our case study, we might set $\beta_{0,0} > \beta_{1,0}$, under the assumption that someone truly in a nonsmoking state would be more likely to report accurately.

Case Study

For the case study, we applied our proposed method to identify environmental, affective, behavioral, social, and baseline factors associated with transitions between smoking states and potentially misreporting smoking behaviors in data collected in the PREVAIL II study. Details of the potential risk factors explored in this analysis are found in the online supplemental materials. For this analysis, data were collected on 198 subjects with a median number of 116 (80–131 IQR) assessments each. Reported momentary smoking, the outcome of interest, is defined as whether or not a subject reported smoking in the 4 hr prior to the current EMA, capturing momentary smoking behaviors during waking hours. However at each EMA, a subject was prompted on their *current* psychological, social, environmental, and behavioral status. Thus, to maintain temporality, we assume that the probability of transition to the current state from the previous state (in the latent space) depends on the covariates observed at the previous state. Similar temporal assumptions have been made previously in smoking behavior research studies (Bolman et al., 2018; Koslovsky et al., 2018; Minami et al., 2014; Shiffman et al., 1996). Under this assumption, regression coefficients are interpreted as log hazard ratios of momentary smoking by the next assessment for a particular risk factor. In this analysis, the initial hidden state represents the first (true) smoking status for subjects after a scheduled quit attempt. Thus, we found the assumption of similar nonsmoking probabilities immediately after a quit attempt reasonably justified. We set $v_k = 5$, which places a 95% prior probability of included regression coefficients between a hazards/odds ratio of .01 and 80. We assumed a 10% prior probability of inclusion for each covariate, parameterized by $a_k = 1$, and $b_k = 9$. This weakly informative prior assumption reflects the exploratory nature of our study aimed at learning potential relations between risk factors and smoking behaviors with little information regarding their occurrence in the presence of other risk factors. For posterior inference, we initiated the MCMC algorithm with the null model (i.e., $\boldsymbol{\beta} = \mathbf{0}$). We ran the MCMC sampler for 80,000 iterations, using the first 25,000 as burn-in and thinning to every 25$25^{th}$ iteration. Trace plots of the parameters' posterior samples indicated good convergence and mixing. Additionally to assess convergence, we ran the model, initiated at the maximum likelihood estimate for regression coefficients using the msm package in R (Jackson, 2011). After burn-in, the correlation for the MPPIs between the two chains was 98.4%, and the multivariate Gelman-Rubin statistic (often referred to as the multivariate potential scale reduction factor) was less than 1.2 for the $\boldsymbol{\beta}$ selected in both models (Gelman & Rubin, 1992), further demonstrating that the MCMC procedure was working properly and the chains converged. The final results were based on the combination of the two chains post-burn-in and thinning, resulting in 4,400 posterior samples for inference. Inclusion in the model for both transition and emission terms was determined using a BFDR threshold of 5%, corresponding to an MPPI $\geq .79$. Goodness-of-fit for the combined chains was assessed using posterior predictive checks, in which we compared replicated data sets from the posterior predictive distribution of the model to the observed data as described in Gelman et al. (2000). Details of the model fit assessment are found in the online supplemental materials. No label switching issues were observed in our case study.

Results of the postquit analysis are found in Table 1. We observed that the contingency management (CM) treatment plan was protective against transitions from nonsmoking to smoking states. Similar results were found in Koslovsky et al. (2018) with a traditional MSM model. We also identified urge to smoke as having a positive relation with transitions from smoking to nonsmoking states. While similar results were found in Koslovsky et al. (2018), these findings may reflect inconsistencies typically found in the relation between urge to smoke and smoking behaviors (Wray et al., 2013). Participants in the present study were provided with nicotine replacement therapy, including nicotine gum, and were instructed to use it when they had an urge to smoke. Nicotine gum is associated with reduction of acute craving in response to exposure to smoking cues (Shiffman et al., 2003), thus it is possible that consumption of nicotine gum may act as a mediator between urge to smoke and smoking behavior. Nicotine gum use was not measured on a momentary basis, but future analyses may provide insight as to why urge to smoke is associated with both transitions to smoking and nonsmoking states. In addition to identifying risk factors associated with transitions between discrete smoking states, our model is also able to identify factors associated with properly reporting their nonsmoking or smoking status (see Table 1). As the accurate reporting of smoking profoundly impacts our understanding of the psychological, social, and environmental contexts that precede high risk moments for smoking lapse, intervention content or decision rules based on false reporting of smoking status may be inadequate or ineffective. This model provides a useful method for accommodating potential reporting bias in situations where remote biochemical verification of smoking status is not possible.

Validation Study

In practice, it is difficult to validate subjects' true smoking behaviors due to limitations in remote biochemical verification technology (e.g., cost and lack of portability). However, in our case study, subjects' smoking behaviors were biochemically validated by collecting their expired carbon monoxide (CO) levels at weekly scheduled in-person follow-ups. While not necessary for the implementation of our proposed method, we used subjects' expired CO levels to assess misreported smoking behaviors and to compare to the misreporting estimated by our proposed approach. Because CO levels tests indicate smoking in the previous 24 hr, only a fraction of the reported smoking behaviors collected via EMAs were available for validation.

For the purposes of validation, smoking status was based on self-report of smoking/abstinence during the past 24 hr on the quit day and during the past 7 days at all follow-ups thereafter. For those who self-reported smoking at each weekly follow-up visit, CO was not further considered even if it was below the threshold (i.e., individuals are not likely to report smoking when they are abstinent). For those who self-reported abstinence, CO had to be 10 ppm or less on the quit day, and 6 ppm or less on subsequent tests to verify abstinence.

In total, 824 CO verifications were performed across the study period for this sample, with an average (*SD*) of 4.20 (1.28) per subject. Only 686 of the 824 CO tests could be used to validate at

Table 1
Case Study Results

Transitions			
Risk Factor	N → S [95% CI]	Risk Factor	S → N [95% CI]
Female	2.387 [2.006, 2.811]	Urge	1.381 [1.206, 1.612]
Non-White	1.406 [1.181, 1.677]	Availability	0.449 [0.357, 0.551]
CM Treatment	0.348 [0.286, 0.419]	Frustrated	0.823 [0.739, 0.908]
Urge	1.685 [1.406, 1.934]	Age	1.401 [1.212, 1.630]
Self-efficacy	0.735 [0.656, 0.830]	HSI	0.701 [0.637, 0.775]
Age	1.343 [1.156, 1.560]		
Emissions			
Risk Factor	Properly report nonsmoking	Risk Factor	Properly report smoking state
Social	0.209 [0.061, 0.513]	Social	2.059 [1.552, 2.861]
Alcohol	0.184 [0.088, 0.383]	Restrict	0.485 [0.339, 0.671]
Availability	0.432 [0.180, 0.827]	Stress	1.303 [1.163, 1.475]
Self-efficacy	1.899 [1.393, 2.601]	Availability	1.872 [1.562, 2.270]
		Self-efficacy	0.735 [0.627, 0.853]

Note. N = nonsmoking; S = smoking; CM = contingency management; HSI = heaviness of smoking index. Estimated hazard ratios and corresponding 95% credible intervals [CI] of selected risk factors for transitioning between N → S and S → N states and covariate specific odds ratios and corresponding 95% CIs of properly reporting nonsmoking and smoking status for the PREVAIL II study postquit date observations.

least one EMA in the previous 24 hr, with a total of 1,831 EMAs collected within the 24-hr windows prior to the CO tests. We define a misreported smoking (nonsmoking) behavior as a passed (failed) CO test, but *at least* one (no) reported smoking behavior in the 24 hr prior to verification. A total of 48 (7.02%) misreported smoking and 117 (17.1%) misreported nonsmoking incidences were observed among all 686 CO verifications, with an average (*SD*) of .25 (.63) and .60 (.89) per person, respectively. For comparison, we estimated each hidden state, h_{ij}, marginally by averaging over their corresponding MCMC samples, with estimates greater than .50 indicating smoking. Using this approach, 314 (2.1%) of the 14,736 estimated nonsmoking hidden states were misreported as smoking states and 851 (16.2%) of the 5,254 estimated smoking states were misreported as nonsmoking states. For the EMAs that fell within the 24-hr CO verification periods, 28 (2.1%) of the 1,348 estimated nonsmoking hidden states were misreported as smoking states, and 66 (13.7%) of the 483 estimated smoking states were misreported as nonsmoking states. For the EMAs that fell outside of the 24-hr CO verification periods, 286 (2.1%) of the 13,388 estimated nonsmoking hidden states were misreported as smoking states, and 785 (16.5%) of the 4,771 estimated smoking states were misreported as nonsmoking states.

Research on the validity of self-reported smoking has demonstrated that self-reports from participants in intervention studies can have lower specificity due to biochemical tests' limited ability to detect very low levels of smoking and less recent smoking (Patrick et al., 1994). Therefore, the small number of nonsmoking states identified by the CO test that were misreported as smoking states may be due to errors in the measurement of CO levels. In the PREVAIL II study, participants in the CM treatment arm had the opportunity to earn gift cards for self-reported and biochemically verified abstinence at each visit. We observed that self-reported smoking status on EMAs occurring within the 24-hr CO verification period had higher concordance with CO results than self-reported status on EMAs outside of the 24-hr verification period. Future research is needed to determine if this is due to participants' awareness of their smoking status being biochemically validated, or if it is reflective of day to day variation in smoking and abstinence states during a quit attempt. In Figures 2 and 3, we present an example of how our model was able to catch a misreported smoking behavior by a subject.

Sensitivity Analysis

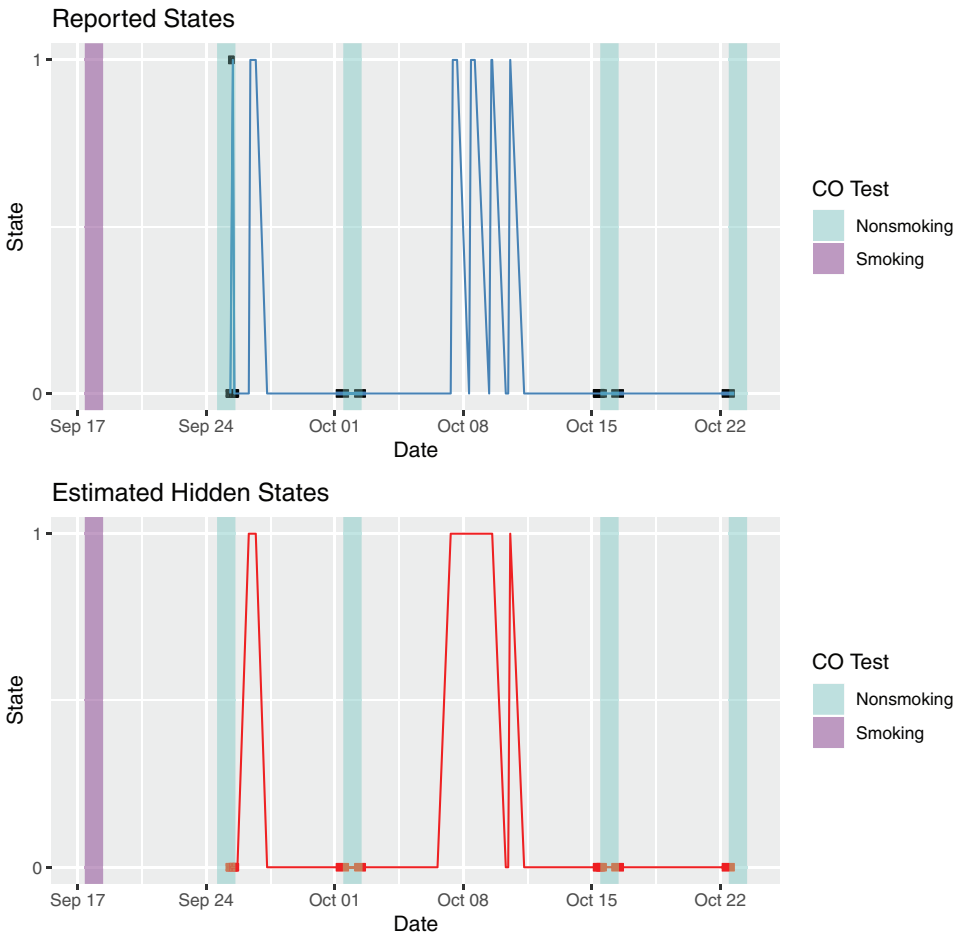
Sensitivity analyses are an integral part of a thorough Bayesian analysis which allow researchers to understand how prior and likelihood specification may influence posterior inference. See van de Schoot et al. (2021) for general guidelines. To investigate the HMM's sensitivity to prior specification, we set each of the hyperparameters to default values and then evaluated the effect of manipulating each term on the results obtained in the case study. For the default parameterization, we used the settings described in the case study. We ran our MCMC algorithm for 80,000 iterations, treating the first 10,000 iterations as burn-in and thinning to every 25^{th} iteration. For each of the transition and emission effects, inclusion in the model was determined using a BFDR threshold of 5%.

Because the true model is never known in practice, we compared the results from the case study with each model parameterization in terms of overall sparsity levels and the overlap in selected covariates. Specifically, we report the total number of terms associated with transitions and emissions as well as the proportion of included and excluded terms in the case study results that were also included or excluded by each parameterization for both transitions (T-IN and T-EX) and emissions (E-IN and E-EX), respectively. Results of the sensitivity analysis are reported in Table 2. We observed little sensitivity to the variance specified for regression coefficients for both transition and emission terms. However, we found some sensitivity to the prior probability of inclusion for transition and emission terms, a potential artifact of the relatively weak associations identified for some of the risk factors. In the following section, we provide recommendations for hyperparameter specifications based on the simulation and sensitivity analysis results.

Additionally to assess the sensitivity of the results to model misspecification, we applied the MSM model with similar hyperparameter settings as the HMM to the case study data. This approach

Figure 2

Reported and Estimated Smoking Behaviors Compared to the CO Verification Results for a Select Subject in the Case Study



Note. CO = carbon monoxide. Upper plot: Blue line represents reported smoking behaviors. Lower plot: Red line indicates estimated hidden behaviors with the proposed model. Shaded windows represent the 24 hr periods prior to CO verification test. See the online article for the color version of this figure.

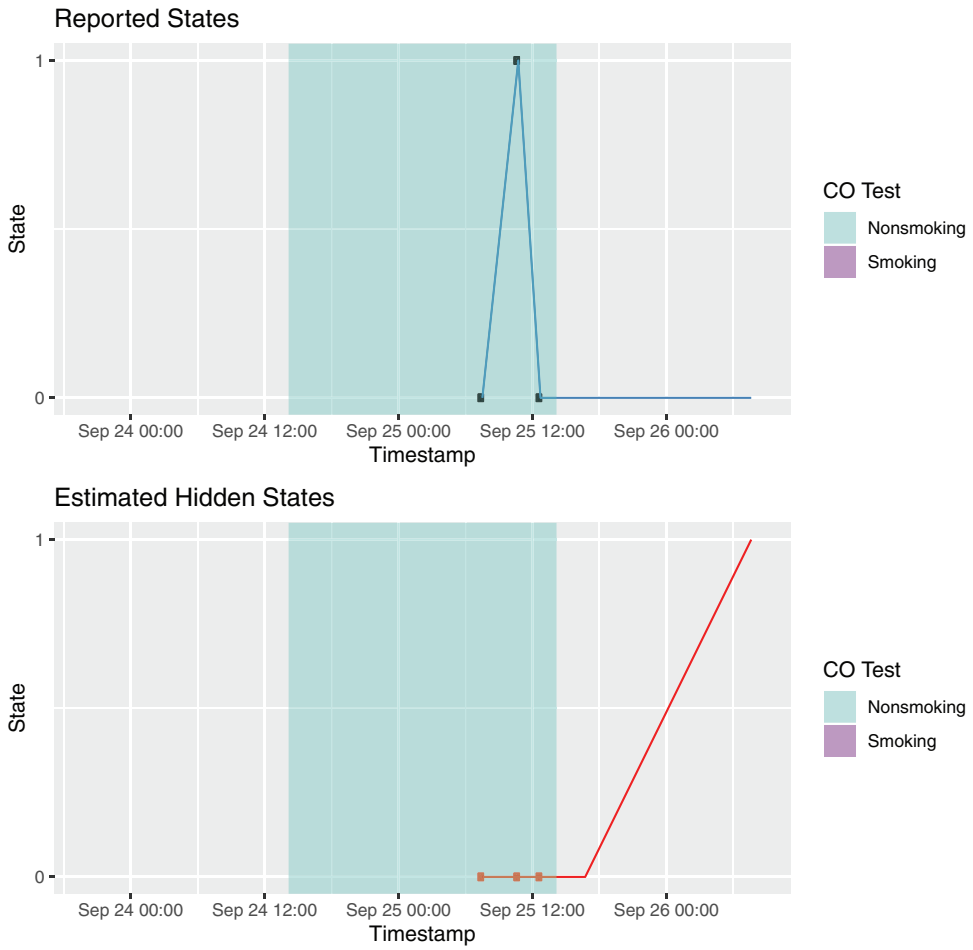
ignores any reporting errors which may produce biased results for estimated hazards ratios for transitions. Thus, we compare the results of the two methods in terms of overlap between identified associations. The MSM model identified a majority of the relations between covariates and transitions between smoking states found with the HMM. However, the MSM model also identified four additional relations for transitions between nonsmoking and smoking states, as well as two additional relations for transitions between smoking and nonsmoking states. In our simulations, we found that the MSM model, in the presence of reporting errors, was prone to higher false positive rates in addition to biased estimates. Thus, it was not surprising that the MSM model potentially overselected.

Simulation Study

We evaluated the variable selection performance of our model on simulated data and compared our method with a continuous-time MSM model in various scenarios, averaging results over 30 replicate data sets for each scenario. Recall that the MSM model is equivalent to removing the emission layer of the HMM and treating the observed states as truth for transitions. In Scenario 1, we mimicked the structure of the case study data, generating data for $\dot{N} = 100$ subjects with $n_i = 31$ assessments, corresponding to $n_i - 1 = 30$ transitions. At each assessment, we simulated $\dot{p} = 30$ potential covariates for the transitions process and $\ddot{p} = 20$ potential covariates for the emission process. In each of the replicate data sets, the first four transition covariates and the first two emission covariates were binary indicators. The rest of the covariates were simulated from a multivariate normal distribution with mean $\mathbf{0}$ and variance Σ, with $\Sigma_{ss} = 1$ and $\Sigma_{st} = 0$. Covariates were standardized prior to analysis. In the true model, we set 20% of the covariates active. Regression coefficients for active terms were randomly set to $\pm\{.7, .85, 1.0\}$, corresponding to odds/hazards ratios ranging from .37 to 2.71. We generated 90% accuracy levels for the observed outcomes, y_{ij}, on average, by setting the intercept terms in the emission models, $\beta_{0,0}$ and

Figure 3

Reported and Estimated Hidden Smoking Behaviors for the Subject in Figure 2 from September 23 at 6:00 p.m. to September 26 at 12:00 p.m



Note. CO = carbon monoxide. Shaded windows represent the 24 hr period prior to CO verification test. Results demonstrate our method's ability to identify discrepancies in reporting and true smoking behaviors, as validated by the expired CO test. See the online article for the color version of this figure.

$\beta_{1,0}$, to logit(.9). Additionally, we investigated 60%, 80%, 95%, and 99% accuracy levels. Moreover, for a baseline comparison, we evaluated the MSM model with 100% accurately observed states. Baseline transitions rates, μ_0 and λ_0 were set to log(.5). Subjects were initialized in a 1 or 0 state with equal probability. In subsequent scenarios, we altered various aspects of Scenario 1 to further investigate the performance of our model. In Scenario 2, we only generated $\dot{N} = 50$ subjects with $n_i = 16$ assessments, corresponding to 15 transitions. In Scenario 3, we increased the potential covariate space with $\dot{p} = \ddot{p} = 100$. In Scenario 4, we reduced the effect sizes to $\pm\{.2, .35, .5\}$. Lastly, in Scenario 5, we investigated the performance of the model with correlated covariates (i.e., $\Sigma_{st} = \omega^{|s-t|}$ where $\omega = .6$).

When running the MCMC algorithm for both HMMs and MSM models, we set the hyperparameters for the prior probability of inclusion θ_k to be noninformative (i.e., $a_k = 1$ and $b_k = 1$, which places equal probability on selection or exclusion for each covariate). We set the regression coefficient prior variance $v_k = 5$ in both models, which assumes a prior probability of included regression coefficients between a hazards/odds ratio of .01 and 80. We used a normal proposal distribution for β_k with mean equal to the current MCMC iteration and variance one. Both the HMMs and MSM models were initiated with all regression coefficients set to zero, except for the intercept terms which were set to one. For the 60% baseline emission probability settings, we observed label switching in some of the replicate data sets. To handle any label switching issues, emission terms were warmstarted with maximum likelihood estimates, obtained using the msm package in R. In each simulation setting, we ran the HMM (MSM) MCMC algorithms for 50,000 (15,000) iterations, treating the first 25,000 (7,500) iterations as burn-in. Inclusion in the model was determined using the median model approach.

Variable selection performance was evaluated via true positive rate (TPR), false positive rate (FPR), and Matthew's correlation coefficient (MCC) for transition and emission effects separately. These metrics are defined as

Table 2
Results of the Sensitivity Analysis for the Case Study

Measure	$a_k = 1$, $b_k = 1$	$a_k = 1$, $b_k = 99$	$v_k = 1$	$v_k = 10$
# Transitions selected	19	7	12	8
T-IN	0.579	1.000	0.917	1.000
T-EX	1.000	0.918	1.000	0.938
# Emissions selected	11	7	9	9
E-IN	0.818	1.000	1.000	1.000
E-EX	1.000	0.953	1.000	1.000

Note. T-IN, T-EX, E-IN, and E-EX represent the proportion of included and excluded terms in the case study results that were also included or excluded by each parameterization for both transitions and emissions, respectively.

$$FPR = \frac{TP}{FN + TP}$$
$$FNR = \frac{FP}{FP + TN}$$
$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}},$$

where *TN*, *TP*, *FN*, and *FP* represent the true negatives, true positives, false negatives, and false positives, respectively. In addition to selection performance, we also evaluated the models' ability to estimate regression coefficients accurately as well as identify discrepancies between hidden and observed states. Specifically, we compared the models in terms of absolute relative bias for active regression coefficients associated with transitions between states, calculated as the absolute value of the difference in estimated and true regression coefficients divided by the true regression coefficient, and the proportion of mismatched hidden and observed states. Hidden states were estimated marginally following Scott (2002).

The variable selection results of our simulation study are presented in Tables 3 and 4. Table 3 presents the results for the transition terms. For all baseline emission probabilities, the HMM outperformed the MSM model in terms of the TPR, FPR, and MCC. We observed a positive relation between the MCC for transition terms and the percentage of accurately reported outcomes. For 60% to 90% baseline emission probabilities, the model obtained around a 10% FPR for transition terms, which decreased to less than 5% above a 95% baseline emission probability. Similarly, we observed MCC values above 70% when applied to data with more than 95% accuracy levels. We found that smaller sample sizes (Scenario 2), larger covariate spaces (Scenario 3), and correlation among the covariates (Scenario 5) reduced the TPR and overall performance of the model for transition terms based on MCC compared with Scenario 1. Both the TPR and FPR decreased with smaller effect sizes (Scenario 4), as expected.

Table 4 contains the results for the emission terms. Here, we found higher TPRs and MCC values for accuracy levels around 80%–90%. The FPR was controlled at $< 4\%$ for all reporting error percentages. We found that the selection performance for emission terms declined above a 90% baseline emission probability. As the proportion of hidden and observed states matching increases, the model is unable to provide accurate results for variable selection in the emission portion of the model due to the reduced variability in the outcomes given the hidden states. Again, we observed poorer selection performance overall for Scenarios 2–5 compared with Scenario 1.

Table 5 contains the results of the bias comparison between the MSM and HMM. Compared with the MSM model, the HMM had lower absolute relative bias for moderate levels of misreporting and relatively equal amounts when there were large amounts of or little misreporting. We observed a sixfold increase in mean absolute relative bias from 100% to 60% baseline emission probabilities using the HMM. Table 6 presents the results for the proportion of matched hidden and estimated states for various baseline emission probabilities. Here, we observed that the proportion of matched states increased with baseline emission probabilities.

Overall, our simulation results show how ignoring potential measurement error can negatively affect variable selection performance and estimation accuracy. Similar to other variable selection methods, we also found limited performance of our model for small sample sizes and correlated covariates. In practice, researchers can control the sparsity level of the model by adjusting the prior probability of inclusion. For example, our simulation study suggests that higher FPRs are obtained with stronger correlations between covariates. To reduce the overall FPR, researchers could impose a smaller prior probability of inclusion to help screen out false signals. Likewise, if the research objective is to generate a larger pool of covariates to validate in future studies, a noninformative prior (which does not induce any sparsity) could be implemented. Our simulation study demonstrates that variable selection and estimation performance declined for both models when applied to data with high levels of misreporting. While these results were not surprising (i.e., garbage in, garbage out), we recommend using other strategies to improve data quality prior to analysis or limiting inference in these settings.

To assess the proposed model's sensitivity to hyperparameter specification, we set each of the hyperparameters to default values and then evaluated the effect of manipulating one term at a time on selection performance. For the default parameterization, we set the prior probability of inclusion hyperparameters $a_k = 1$ and $b_k = 1$ and the regression coefficient prior variance to $v_k = 5$, similar to our simulation study. Additionally, we assessed the sensitivity of the selection results to the assumption of similar initial state probabilities across subjects. To evaluate this, we applied the proposed model to data simulated similarly to Scenario 1 with subject-specific initial hidden state probabilities.

We evaluated the model's sensitivity to prior settings using data generated in the simulation study with a baseline emission probability of 90%. We ran the MCMC algorithms for 50,000 iterations, using the first 25,000 iterations as burn-in and the other 25,000 iterations for inference. To initialize each model, we set all regression terms to be excluded except for the intercept terms, similar to the simulation study.

Results of the sensitivity analysis are presented in Table 7. Holding the prior variance $v_k = 5$ constant, we found that decreasing ($b_k = 4$, 9, and 99) the prior probability of inclusion decreased the average number of terms selected, as expected. As mentioned previously, the prior sparsity level selected by the researcher should accurately reflect the research objectives and ideally be determined prior to analysis. Holding $b_k = 1$, we observed little variation in selection performance across different values of the slab prior, v_k. Specifically, the TPR, FPR, and MCC hovered around 60%, 10%, and 50% for the transition terms, respectively.

Table 3
Simulation Results for Comparison of Variable Selection Performance on Transition Terms from 0 → 1 (β_λ) and 1 → 0 (β_μ) Using HMMs and MSM Models

		60%		80%		90%		95%		99%		100%	
Transitions	Measure	HMM	MSM	HMM	MSM	HMM	MSM	HMM	MSM	HMM	MSM	HMM	MSM
0 → 1	TPR	0.205	0.161	0.400	0.444	0.594	0.522	0.694	0.644	0.906	0.856	0.978	0.967
	FPR	0.090	0.118	0.121	0.167	0.113	0.168	0.033	0.111	0.004	0.036	0.006	0.007
	MCC	0.137	0.044	0.293	0.270	0.476	0.343	0.717	0.536	0.930	0.829	0.974	0.963
1 → 0	TPR	0.228	0.244	0.461	0.411	0.578	0.544	0.733	0.678	0.889	0.839	0.956	0.944
	FPR	0.111	0.146	0.125	0.182	0.099	0.194	0.039	0.125	0.004	0.024	0.004	0.008
	MCC	0.127	0.113	0.350	0.224	0.498	0.329	0.739	0.536	0.920	0.843	0.962	0.946

		Scenario 2		Scenario 3		Scenario 4		Scenario 5	
Transitions	Measure	HMM	MSM	HMM	MSM	HMM	MSM	HMM	MSM
0 → 1	TPR	0.322	0.328	0.339	0.406	0.478	0.422	0.483	0.472
	FPR	0.058	0.136	0.069	0.126	0.025	0.058	0.103	0.142
	MCC	0.339	0.208	0.239	0.195	0.572	0.437	0.404	0.333
1 → 0	TPR	0.306	0.339	0.411	0.467	0.506	0.411	0.533	0.489
	FPR	0.083	0.140	0.062	0.128	0.036	0.079	0.119	0.139
	MCC	0.270	0.216	0.309	0.228	0.562	0.380	0.413	0.359

Note. HMM = hidden Markov model; MSM = multistate Markov model; TPR = true positive rate; FPR = false positive rate; MCC = Matthew's correlation coefficient. Results are presented for Scenario 1 for each baseline emission probability (60%, 80%, 90%, 95%, 99% and 100%) and for Scenarios 2–5 averaged across 30 replicate datasets.

Additionally, the TPR, FPR, and MCC hovered around 80%, 5%, and 80% for the emission terms, respectively. Thus, we recommend a moderately diffuse prior for regression coefficients in application. Lastly, we observed little to no impact of the equal initial state probability assumption on selection performance.

Conclusions

In this article, we have developed a fully Bayesian variable selection approach for continuous-time hidden Markov models. Our method treats reported outcomes as emissions from a subject's true, latent state. By employing spike-and-slab variable selection priors, our method is able to identify risk factors for subjects transitioning between discrete latent states as well as risk factors potentially associated with them misreporting their true state. Our method takes advantage of data augmentation techniques for posterior inference to reduce computation time without sacrificing model interpretation. We have applied our method to smartphone-based ecological momentary assessment data collected in a randomized controlled trial designed to investigate the efficacy of a contingency management intervention for smoking cessation in a population of socioeconomically disadvantaged smokers. Our results support prior research and demonstrate the potential effectiveness of financial incentives (contingency management) on short-term smoking cessation outcomes (Notley et al., 2019). In addition, we have identified several key variables associated with transitions to smoking states during a quit attempt, including momentary factors such as interacting with another smoker, urge to smoke, and self-efficacy that may be important targets for future interventions. These findings are consistent with other studies that have used EMAs to explore the psychological, social, and environmental factors associated with smoking during a quit attempt. For example, multiple studies have identified smoking urge (Chandra et al., 2011; Shiffman et al., 2009), proximity to other smokers (Piasecki et al., 2014; Shiffman et al., 2007), and low motivation to quit (Businelle et al., 2014; Hébert et al., 2021) as predictors of smoking lapse. The method presented here has unique implications for intervention development in that we examine not only

Table 4
Simulation Results for Variable Selection Performance on Emission Terms

Emissions	Measure	60%	80%	90%	95%	99%	S2	S3	S4	S5
β_0	TPR	0.533	0.792	0.733	0.617	0.167	0.250	0.325	0.142	0.692
	FPR	0.042	0.017	0.021	0.010	0.008	0.023	0.053	0.004	0.025
	MCC	0.598	0.828	0.781	0.748	0.440	0.398	0.255	0.500	0.750
β_1	TPR	0.683	0.917	0.792	0.722	0.200	0.258	0.442	0.158	0.758
	FPR	0.050	0.010	0.031	0.010	0.010	0.033	0.053	0.015	0.027
	MCC	0.681	0.925	0.811	0.790	0.481	0.378	0.334	0.405	0.789

Note. TPR = true positive rate; FPR = false positive rate; MCC = Matthew's correlation coefficient. Results are presented for Scenario 1 using HMMs for each baseline emission probability (60%, 80%, 90%, 95%, 99% and 100%) and Scenarios 2–5 (S2–S5) averaged across 30 replicate datasets. Emission coefficients β_0 and β_1 correspond to covariates associated with properly reporting 0 and 1 states, respectively.

Table 5
Results of the Absolute Relative Bias Assessment

Model	60%	80%	90%	95%	99%	100%
HMM	0.86	0.70	0.55	0.31	0.16	0.15
MSM	0.83	0.77	0.66	0.50	0.22	0.14

Note. HMM = hidden Markov model; MSM = multistate Markov model. Results are averaged over active transition terms across 30 simulated datasets for each baseline emission probability (60%, 80%, 90%, 95%, 99% and 100%) using the proposed MSM and HMMs.

what factors are associated with lapse, but what factors are positively associated with transitioning to a nonsmoking state. Therefore, interventions that seek to prevent factors associated with lapse, as well as those that actively reinforce factors associated with transitioning back to a nonsmoking state may improve the likelihood of a successful smoking cessation attempt.

EMA is limited by self-report, and the potential for response fatigue and misreporting, whether intentional or not, can significantly impact prediction accuracy. In order to improve the delivery of future mobile health interventions, such as just-in-time adaptive interventions, it is important to know when and how to deliver intervention content to maximize impact. We do not suggest or advocate any particular intervention approach but hope that future research might be able to investigate how to minimize participant burden and utilize objective data to prevent the delivery of intervention content when it is not needed. In addition to accounting for potential misreporting, our method also suggests covariates that may be associated with reporting errors. These insights may help researchers assess data quality and design more effective data collection strategies. While these results highlight the importance of understanding risk factors' relations with smoking behaviors when designing tailored intervention strategies, we recommend using our method for hypothesis generation in practice and conducting confirmatory studies before generalizing results. Further, the associations identified in the case study do not imply causal relations. Because momentary smoking was defined as any reported smoking behaviors in the 4 hr prior to the EMA, the first assessment after waking is potentially subject to a ceiling effect.

Table 6
Simulation Results for the Proportion of Matched Estimated and Hidden States

Emission probability	Proportion of matched states
60%	0.636 (0.120)
80%	0.835 (0.029)
90%	0.890 (0.014)
95%	0.938 (0.012)
99%	0.982 (0.004)

Note. Results are reported as *M*(*SD*) across each simulated baseline emission probability (60%, 80%, 90%, 95%, and 99%).

However, our objective when constructing the outcome was to capture whether or not a subject had smoked within nonoverlapping epochs during waking hours, following the assumptions of our model. Thus, the first 4-hr window captures any morning smoking behaviors if they occurred. Through simulations, we have demonstrated how our model improves variable selection performance and estimation accuracy compared to a traditional MSM model at various levels of reporting errors. As with any analysis, accurate model inference depends on appropriate model specification. In practice, we encourage thorough investigation of posterior convergence, goodness-of-fit assessment, and sensitivity analyses prior to drawing inference on the results generated from our model (i.e., identified misreported outcomes and influential covariates).

Even though we have developed our approach for mHealth data, the proposed model is applicable to other longitudinal settings in which transitions between two-discrete states are measured potentially with error. We provide code to simulate data similar to our simulation study and implement our proposed method in our contributed R package, HMMbvs, available at https://github.com/mliang4/HMMbvs. While developed for two-state continuous HMMs, extensions to three and four state are readily available because closed-form solutions for transition probabilities exist (Li & Chan, 2006). For larger state spaces, approximation methods for solving the ordinary differential equations for transition probabilities would be required. By using a full MCMC approach, our method provides inherent estimates for model uncertainty, compared with alternative variable selection methods

Table 7
Sensitivity Results on Simulated Data

		$v_k = 5$			$b_k = 1$		
Selection measure	Measure	$b_k = 4$	$b_k = 9$	$b_k = 99$	$v_k = 1$	$v_k = 10$	Random π_1
Mean # of selected terms		17.7	15.9	13.0	27.4	22.1	23.8
Transition $0 \rightarrow 1$ (β_λ)	TPR	0.494	0.422	0.378	0.578	0.589	0.572
	FPR	0.072	0.071	0.069	0.107	0.101	0.093
	MCC	0.458	0.406	0.368	0.474	0.495	0.489
Transition $1 \rightarrow 0$ (β_μ)	TPR	0.522	0.450	0.417	0.622	0.556	0.578
	FPR	0.076	0.088	0.078	0.119	0.086	0.104
	MCC	0.490	0.406	0.385	0.498	0.489	0.482
Properly report 0 state (β_0)	TPR	0.683	0.650	0.525	0.767	0.700	0.808
	FPR	0.010	0.006	0.000	0.029	0.019	0.025
	MCC	0.780	0.769	0.706	0.779	0.754	0.822
Properly report 1 state (β_1)	TPR	0.750	0.717	0.542	0.875	0.725	0.850
	FPR	0.010	0.013	0.004	0.033	0.013	0.015
	MCC	0.825	0.783	0.735	0.854	0.818	0.869

Note. TPR = true positive rate; FPR = false positive rate; MCC = Matthew's correlation coefficient. Sensitivity results for the specification of the slab variance hyperparameters, prior probability of inclusion, and random initial probability π_1.

for MSM models which use optimization-based estimation techniques (Koslovsky et al., 2018). However, full inference comes at a price computationally. In larger model spaces, the MCMC chains may require more memory than available. To help manage memory storage, the amount of thinning can be increased. Additionally, variational alternatives may provide reasonably accurate model estimates in a fraction of the time. Recently, Koslovsky et al. (2020) developed a Bayesian, semiparametric logistic regression model which performs variable selection for varying-coefficient terms and random effects. Future work could introduce varying-coefficient terms and random effects into the HMM framework to investigate how a covariate may vary as a function of another covariate as well as how effects may vary across subjects.

References

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press Publications.

Barbieri, M. M., & Berger, J. O. (2004). Optimal predictive model selection. *The Annals of Statistics*, *32*(3), 870–897. https://doi.org/10.1214/009053604000000238

Benoit, J. S., Chan, W., Luo, S., Yeh, H.-W., & Doody, R. (2016). A hidden Markov model approach to analyze longitudinal ternary outcomes when some observed states are possibly misclassified. *Statistics in Medicine*, *35*(9), 1549–1557. https://doi.org/10.1002/sim.6861

Benowitz, N. L., Bernert, J. T., Foulds, J., Hecht, S. S., Jacob, P., Jarvis, M. J., Joseph, A., Oncken, C., & Piper, M. E. (2020). Biochemical verification of tobacco use and abstinence: 2019 update. *Nicotine and Tobacco Research*, *22*(7), 1086–1097. https://doi.org/10.1093/ntr/ntz132

Bentley, K. H., Kleiman, E. M., Elliott, G., Huffman, J. C., & Nock, M. K. (2019). Real-time monitoring technology in single-case experimental design research: Opportunities and challenges. *Behaviour Research and Therapy*, *117*, 87–96. https://doi.org/10.1016/j.brat.2018.11.017

Bertz, J. W., Epstein, D. H., & Preston, K. L. (2018). Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance-use research. *Addictive Behaviors*, *83*, 5–17. https://doi.org/10.1016/j.addbeh.2017.11.027

Blaauw, F. J., Schenk, H. M., Jeronimus, B. F., van der Krieke, L., de Jonge, P., Aiello, M., & Emerencia, A. C. (2016). Let's get physiqual–An intuitive and generic method to combine sensor technology with ecological momentary assessments. *Journal of Biomedical Informatics*, *63*, 141–149. https://doi.org/10.1016/j.jbi.2016.08.001

Bolman, C., Verboon, P., Thewissen, V., Boonen, V., Soons, K., & Jacobs, N. (2018). Predicting smoking lapses in the first week of quitting: an ecological momentary assessment study. *Journal of Addiction Medicine*, *12*(1), 65–71. https://doi.org/10.1097/ADM.0000000000000365

Bond, J. C., Greenfield, T. K., Patterson, D., & Kerr, W. C. (2014). Adjustments for drink size and ethanol content: New results from a self-report diary and transdermal sensor validation study. *Alcoholism: Clinical and Experimental Research*, *38*(12), 3060–3067. https://doi.org/10.1111/acer.12589

Brown, P. J., Vannucci, M., & Fearn, T. (1998). Multivariate Bayesian variable selection and prediction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(3), 627–641. https://doi.org/10.1111/1467-9868.00144

Bureau, A., Shiboski, S., & Hughes, J. P. (2003). Applications of continuous time hidden Markov models to the study of misclassified disease outcomes. *Statistics in Medicine*, *22*(3), 441–462. https://doi.org/10.1002/sim.1270

Businelle, M. S., Ma, P., Kendzor, D. E., Frank, S. G., Vidrine, D. J., & Wetter, D. W. (2016). An ecological momentary intervention for smoking cessation: Evaluation of feasibility and effectiveness. *Journal of Medical Internet Research*, *18*(12), e321. https://doi.org/10.2196/jmir.6058

Businelle, M. S., Ma, P., Kendzor, D. E., Reitzel, L. R., Chen, M., Lam, C. Y., Bernstein, I., & Wetter, D. W. (2014). Predicting quit attempts among homeless smokers seeking cessation treatment: an ecological momentary assessment study. *Nicotine & Tobacco Research*, *16*(10), 1371–1378. https://doi.org/10.1093/ntr/ntu088

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: A modern perspective*. CRC press.

Chandra, S., Scharf, D., & Shiffman, S. (2011). Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug and Alcohol Dependence*, *117*(2–3), 118–125. https://doi.org/10.1016/j.drugalcdep.2010.12.027

Cox, D. R., & Miller, H. D. (2017). *The theory of stochastic processes*. Routledge.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Gelman, A., Goegebeur, Y., Tuerlinckx, F., & Van Mechelen, I. (2000). Diagnostic checks for discrete data regression models using posterior predictive simulations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *49*(2), 247–268.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

George, E. I., & McCulloch, R. E. (1997). Approaches for Bayesian variable selection. *Statistica Sinica*, *7*(2), 339–373.

Gustafson, P. (2003). *Measurement error and misclassification in statistics and epidemiology: Impacts and Bayesian adjustments*. CRC Press.

Hébert, E. T., Ra, C. K., Alexander, A. C., Helt, A., Moisiuc, R., Kendzor, D. E., Vidrine, D. J., Funk-Lawler, R. K., & Businelle, M. S. (2020). A mobile just-in-time adaptive intervention for smoking cessation: Pilot randomized controlled trial. *Journal of Medical Internet Research*, *22*(3), e16907. https://doi.org/10.2196/16907

Hébert, E. T., Suchting, R., Ra, C. K., Alexander, A. C., Kendzor, D. E., Vidrine, D. J., & Businelle, M. S. (2021). Predicting the first smoking lapse during a quit attempt: A machine learning approach. *Drug and Alcohol Dependence*, *218*, 108340.

Herbec, A., Parker, E., Ubhi, H. K., Raupach, T., & West, R. (2020). Decrease in resting heart rate measured using smartphone apps to verify abstinence from smoking: An exploratory study. *Nicotine & Tobacco Research*, *22*(8), 1424–1427. https://doi.org/10.1093/ntr/ntaa021

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, *15*(Pt. 1), 1–39. https://doi.org/10.1348/135910709X466063

Jackson, C. H. (2011). Multi-state models for panel data: The msm package for R. *Journal of Statistical Software*, *38*(8), 1–29. https://doi.org/10.18637/jss.v038.i08

Jasra, A., Holmes, C. C., & Stephens, D. A. (2005). Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. *Statistical Science*, *20*(1), 50–67. https://doi.org/10.1214/088342305000000016

Jones, R. H., Xu, S., & Grunwald, G. K. (2006). Continuous time Markov models for binary longitudinal data. *Biometrical Journal*, *48*(3), 411–419. https://doi.org/10.1002/bimj.200510224

Kalbfleisch, J., & Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, *80*(392), 863–871. https://doi.org/10.1080/01621459.1985.10478195

Kang, K., Song, X., Hu, X. J., & Zhu, H. (2019). Bayesian adaptive group lasso with semiparametric hidden Markov models. *Statistics in Medicine*, *38*(9), 1634–1650. https://doi.org/10.1002/sim.8051

Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. *Developmental Psychology*, *44*(2), 457–467. https://doi.org/10.1037/0012-1649.44.2.457

Kendzor, D. E., Businelle, M. S., Waring, J. J., Mathews, A. J., Geller, D. W., Barton, J. M., Alexander, A. C., Hébert, E. T., Ra, C. K., & Vidrine, D. J. (2020). Automated mobile delivery of financial incentives for smoking cessation among socioeconomically disadvantaged adults: Feasibility study. *JMIR mHealth and uHealth*, *8*(4), e15960. https://doi.org/10.2196/15960

Koslovsky, M. D., Hébert, E. T., Businelle, M. S., & Vannucci, M. (2020). A Bayesian time-varying effect model for behavioral mHealth data. *The Annals of Applied Statistics*, *14*(4), 1878–1902.

Koslovsky, M. D., Swartz, M. D., Chan, W., Leon-Novelo, L., Wilkinson, A. V., Kendzor, D. E., & Businelle, M. S. (2018). Bayesian variable selection for multistate Markov models with interval-censored data in an ecological momentary assessment study of smoking cessation. *Biometrics*, *74*(2), 636–644. https://doi.org/10.1111/biom.12792

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W. T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebvre, R. C., Mohr, D. C., Murphy, S. A., Quinn, C., Shusterman, V., & Swendeman, D. (2013). Mobile health technology evaluation: The mHealth evidence workshop. *American Journal of Preventive Medicine*, *45*(2), 228–236. https://doi.org/10.1016/j.amepre.2013.03.017

Li, Y.-P., & Chan, W. (2006). Analysis of longitudinal multinomial outcome data. *Biometrical Journal*, *48*(2), 319–326. https://doi.org/10.1002/bimj.200510187

Li, T., & Vuong, Q. (1998). Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis*, *65*(2), 139–165. https://doi.org/10.1006/jmva.1998.1741

Ma, J., Chan, W., & Tilley, B. C. (2018). Continuous time Markov chain approaches for analyzing transtheoretical models of health behavioral change: A case study and comparison of model estimations. *Statistical Methods in Medical Research*, *27*(2), 593–607. https://doi.org/10.1177/0962280216639859

Ma, J., Chan, W., Tsai, C.-L., Xiong, M., & Tilley, B. C. (2015). Analysis of transtheoretical model of health behavioral changes in a nutrition intervention study—a continuous time Markov chain model with Bayesian approach. *Statistics in Medicine*, *34*(27), 3577–3589. https://doi.org/10.1002/sim.6571

McClure, E. A., Tomko, R. L., Carpenter, M. J., Treiber, F. A., & Gray, K. M. (2018). Acceptability and compliance with a remote monitoring system to track smoking and abstinence among young smokers. *The American Journal of Drug and Alcohol Abuse*, *44*(5), 561–570. https://doi.org/10.1080/00952990.2018.1467431

Minami, H., Yeh, V. M., Bold, K. W., Chapman, G. B., & McCarthy, D. E. (2014). Relations among affect, abstinence motivation and confidence, and daily smoking lapse risk. *Psychology of Addictive Behaviors*, *28*(2), 376–388. https://doi.org/10.1037/a0034445

Mitchell, H. B. (2007). *Multi-sensor data fusion: An introduction*. Springer Science & Business Media.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Nelson, B. W., & Allen, N. B. (2018). Extending the passive-sensing toolbox: Using smart-home technology in psychological science. *Perspectives on Psychological Science*, *13*(6), 718–733. https://doi.org/10.1177/1745691618776008

Newton, M. A., Noueiry, A., Sarkar, D., & Ahlquist, P. (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics*, *5*(2), 155–176. https://doi.org/10.1093/biostatistics/5.2.155

Notley, C., Gentry, S., Livingstone-Banks, J., Bauld, L., Perera, R., & Hartmann-Boyce, J. (2019). Incentives for smoking cessation. *Cochrane Database of Systematic Reviews*, *7*, CD004307. https://doi.org/10.1002/14651858.CD004307.pub6

Papastamoulis, P. (2016). label.switching: An R package for dealing with the label switching problem in MCMC outputs. *Journal of Statistical Software*. Advance online publication. https://doi.org/10.18637/jss.v069.c01

Patrick, D. L., Cheadle, A., Thompson, D. C., Diehr, P., Koepsell, T., & Kinne, S. (1994). The validity of self-reported smoking: A review and meta-analysis. *American Journal of Public Health*, *84*(7), 1086–1093. https://doi.org/10.2105/ajph.84.7.1086

Peng, H.-L., Aschenbrenner, A., von Sternberg, K., Mullen, P. D., & Chan, W. (2019). A continuous-time Markov chain approach with the analytic likelihood in studies of behavioral changes. *Communications in Statistics-Theory and Methods*, *48*(23), 5756–5765. https://doi.org/10.1080/03610926.2018.1520886

Piasecki, T. M., Trela, C. J., Hedeker, D., & Mermelstein, R. J. (2014). Smoking antecedents: Separating between-and within-person effects of tobacco dependence in a multiwave ecological momentary assessment investigation of adolescent smoking. *Nicotine & Tobacco Research*, *16* (Suppl. 2), S119–S126. https://doi.org/10.1093/ntr/ntt132

Pinsky, M., & Karlin, S. (2010). *An introduction to stochastic modeling*. Academic press.

Polson, N. G., Scott, J. G., & Windle, J. (2013). Bayesian inference for logistic models using Pólya–Gamma latent variables. *Journal of the American Statistical Association*, *108*(504), 1339–1349. https://doi.org/10.1080/01621459.2013.829001

Rashid, N., Sun, W., & Ibrahim, J. G. (2014). Some statistical strategies for DAE-seq data analysis: Variable selection and modeling dependencies among observations. *Journal of the American Statistical Association*, *109*(505), 78–94. https://doi.org/10.1080/01621459.2013.869222

Redner, R. A., & Walker, H. F. (1984). Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review*, *26*(2), 195–239. https://doi.org/10.1137/1026034

Rehg, J. M., Murphy, S. A., & Kumar, S. (2017). *Mobile health*. Springer.

Reulen, H., & Kneib, T. (2016). Boosting multi-state models. *Lifetime Data Analysis*, *22*(2), 241–262. https://doi.org/10.1007/s10985-015-9329-9

Savitsky, T., Vannucci, M., & Sha, N. (2011). Variable selection for nonparametric Gaussian process priors: Models and computational strategies. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *26*(1), 130–149. https://doi.org/10.1214/11-STS354

Schukat, M., McCaldin, D., Wang, K., Schreier, G., Lovell, N. H., Marschollek, M., & Redmond, S. J. (2016). Unintended consequences of wearable sensor use in healthcare. *Yearbook of Medical Informatics*, *25*(1), 73–86. https://doi.org/10.15265/IY-2016-025

Scott, S. L. (2002). Bayesian methods for hidden Markov models: Recursive computing in the 21st century. *Journal of the American Statistical Association*, *97*(457), 337–351. https://doi.org/10.1198/016214502753479464

Sennhenn-Reulen, H., & Kneib, T. (2016). Structured fusion lasso penalized multi-state models. *Statistics in Medicine*, *35*(25), 4637–4659. https://doi.org/10.1002/sim.7017

Shiffman, S., Balabanis, M. H., Gwaltney, C. J., Paty, J. A., Gnys, M., Kassel, J. D., Hickcox, M., & Paton, S. M. (2007). Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. *Drug and Alcohol Dependence*, *91*(2-3), 159–168. https://doi.org/10.1016/j.drugalcdep.2007.05.017

Shiffman, S., Gwaltney, C. J., Balabanis, M. H., Liu, K. S., Paty, J. A., Kassel, J. D., Hickcox, M., & Gnys, M. (2009). Immediate antecedents of cigarette smoking: an analysis from ecological momentary assessment. *Addictive Behaviors: New Readings on Etiology, Prevention, and Treatment*, *34*(6-7), 514–519. https://doi.org/10.1016/j.addbeh.2009.01.004

Shiffman, S., Paty, J. A., Gnys, M., Kassel, J. A., & Hickcox, M. (1996). First lapses to smoking: within-subjects analysis of real-time reports. *Journal of Consulting and Clinical Psychology*, *64*(2), 366–379. https://doi.org/10.1037/0022-006X.64.2.366

Shiffman, S., Shadel, W. G., Niaura, R., Khayrallah, M. A., Jorenby, D. E., Ryan, C. F., & Ferguson, C. L. (2003). Efficacy of acute administration of nicotine gum in relief of cue-provoked cigarette craving. *Psychopharmacology*, *166*(4), 343–350. https://doi.org/10.1007/s00213-002-1338-1

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Spezia, L. (2020). Bayesian variable selection in non-homogeneous hidden Markov models through an evolutionary Monte Carlo method. *Computational Statistics & Data Analysis*, *143*, 106840. https://doi.org/10.1016/j.csda.2019.106840

Stephens, M. (2000). Dealing with label switching in mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *62*(4), 795–809. https://doi.org/10.1111/1467-9868.00265

Stingo, F. C., Chen, Y. A., Vannucci, M., Barrier, M., & Mirkes, P. E. (2010). A Bayesian graphical modeling approach to microrna regulatory network inference. *The Annals of Applied Statistics*, *4*(4), 2024–2048. https://doi.org/10.1214/10-AOAS360

Stingo, F. C., & Vannucci, M. (2011). Variable selection for discriminant analysis with Markov random field priors for the analysis of microarray data. *Bioinformatics*, *27*(4), 495–501. https://doi.org/10.1093/bioinformatics/btq690

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

Van Erp, S., Oberski, D. L., & Mulder, J. (2019). Shrinkage priors for Bayesian penalized regression. *Journal of Mathematical Psychology*, *89*, 31–50. https://doi.org/10.1016/j.jmp.2018.12.004

Visser, I. (2011). Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *Journal of Mathematical Psychology*, *55*(6), 403–415. https://doi.org/10.1016/j.jmp.2011.08.002

Visser, I., Raijmakers, M. E., & Molenaar, P. (2002). Fitting hidden Markov models to psychological data. *Scientific Programming*, *10*(3), 185–199. https://doi.org/10.1155/2002/874560

Walls, T. A. (2013). Intensive longitudinal data. *The Oxford Handbook of Quantitative Methods: Statistical Analysis*, *2*, 432–440.

Wray, J. M., Gass, J. C., & Tiffany, S. T. (2013). A systematic review of the relationships between craving and smoking cessation. *Nicotine & Tobacco Research*, *15*(7), 1167–1182. https://doi.org/10.1093/ntr/nts268

Yen, A. M.-F., & Chen, H.-H. (2018). Bayesian measurement-error-driven hidden Markov regression model for calibrating the effect of covariates on multistate outcomes: Application to androgenetic alopecia. *Statistics in Medicine*, *37*(21), 3125–3146. https://doi.org/10.1002/sim.7813

Zhou, J., Song, X., & Sun, L. (2020). Continuous time hidden Markov model for longitudinal data. *Journal of Multivariate Analysis*, *179*, 104646. https://doi.org/10.1016/j.jmva.2020.104646

Received December 10, 2020
Revision received July 22, 2021
Accepted August 4, 2021 ■