Identification of physiological shock in intensive care units via Bayesian regime switching models

Emmett B. Kendall[1], Jonathan P. Williams[2], Curtis B. Storlie[3], Misty A. Radosevich[3], Erica D. Wittwer[3], and Matthew A. Warner[3]

[1]Department of Mathematical Sciences, The University of Texas at Dallas
[2]Department of Statistics, North Carolina State University
[3]Department of Quantitative Health Sciences, Mayo Clinic

March 18, 2026

Abstract

Detection of occult hemorrhage (i.e., internal bleeding) in patients in intensive care units (ICUs) can pose significant challenges for critical care workers. Because blood loss may not always be clinically apparent, clinicians rely on monitoring vital signs for specific trends indicative of a hemorrhage event. The inherent difficulties of diagnosing such an event can lead to late intervention by clinicians which has catastrophic consequences. Therefore, a methodology for early detection of hemorrhage has wide utility. We develop a Bayesian regime switching model (RSM) that analyzes trends in patients' vitals and labs to provide a probabilistic assessment of the underlying physiological state that a patient is in at any given time. This article is motivated by a comprehensive dataset we curated from Mayo Clinic of 33,924 real ICU patient encounters. Longitudinal response measurements are modeled as a vector autoregressive process conditional on all latent states up to the current time point, and the latent states follow a Markov process. We present a novel Bayesian sampling routine to learn the posterior probability distribution of the latent physiological states, as well as develop an approach to account for pre-ICU-admission physiological changes. A simulation and real case study illustrate the effectiveness of our approach.

Keywords: Biomedical data, electronic health records, hidden Markov model, hierarchical Bayes, state-space model

1 Introduction

Hemorrhage, especially occult blood loss, is a serious and potentially life threatening complication that is known to be difficult to diagnose. Notably, 20-40% of hospital patients die due to injury-related-hemorrhage that was *preventable* if the internal bleeding had been recognized earlier (Holcomb et al. 2015). Additionally, in those with trauma-related hemorrhage, 40% of preventable deaths are related to inadequate hemorrhage recognition or control (Holcomb et al. 2015, Stensballe et al. 2017). Patients that have experienced serious trauma and/or hypovolemic shock see an increased risk of more severe internal bleeding events (Moore et al. 2021). There are many difficulties in the diagnosis of bleeding, and part of the problem is the wide variation in what precisely defines a "major bleed" coupled with the fact that many other medical ailments can disguise the canonical behavior of hemorrhage (Maier et al. 2024). This difficulty in detection can ultimately mean delayed diagnoses and care, which can lead to physiological shock and possibly death.

The motivation for this work comes from a comprehensive dataset comprising of 33,924 patient encounters that our team of researchers has curated from Mayo Clinic's ICUs. Clinicians at Mayo Clinic are interested in a data-driven and model-based approach to reduce delay in the diagnosis of internal bleeding. The model developed should, (1) account for inter-individual differences in physiological response to hemorrhage, (2) be robust to missingness in vital and lab measurements, (3) incorporate medication information, and (4) account for pre-ICU-admission physiological changes.

The major contribution of our work is an innovative and in-depth case study about the performance of using a Bayesian semi-supervised RSM to detect internal bleeding. Many challenges exist for this application, requiring novel statistical innovation. First, we develop a unique state-sampling routine to efficiently estimate the discrete posterior distribution of

the latent physiological state for every time point for each patient encounter, both defining and estimating labels for patients' physiological conditions. Second, we fit a parametric model such that the results are interpretable to a clinician as opposed to more common, "black-box" machine learning (ML) approaches. Third, the mean structure for our Bayesian RSM is able to account for physiological changes prior to a patient's ICU admission. Lastly, we test the efficacy of our approach on a small test set of patient encounters that have been manually clinically annotated, to determine the accuracy of model predictions on real data.

With advancements in computing technology and statistical learning methodology, the development of RSMs (or state-space models) for applications to the biological and medical sciences has grown markedly (Kalbfleisch & Lawless 1985, Satten & Longini 1996, Bureau et al. 2003, Jackson et al. 2003, Scott et al. 2005, Altman 2007, Shirley et al. 2010, Zhao et al. 2016, Langrock et al. 2018, Li et al. 2019, Williams et al. 2020, Sidrow et al. 2022). Many of these applications are with hidden Markov models (HMMs), perhaps the most fundamental example of an RSM. An HMM, and RSM more broadly, is used to model two simultaneous stochastic processes: an observed response process and a hidden state process. For a thorough review of HMMs, see Rabiner (1989); for more recent applications of HMMs, see Storlie et al. (2014), Kendall et al. (2025), Volpe et al. (2025). The full potential of the RSM framework for biomedical research, however, remains to be realized, and this is particularly true for hierarchical Bayesian constructions of RSMs with applications for patient monitoring. Statistical learning techniques for detecting adverse physiological events are inherently hard to train because data labels are often difficult to characterize or simply unavailable. Dealing with a lack of gold-standard-labeled training data is a common challenge for RSM applications. This challenge is addressed in Trabelsi et al. (2013) where, as they describe, an HMM can be trained on data to learn latent states in the absence

of annotated data, in an unsupervised fashion, assuming the number of latent states to learn is known. In this case, the HMM acts as a classification algorithm and takes into account time-series regime changes to characterize each latent state. In order to better control the model complexity through prior density specifications, we note the utilization of unsupervised Bayesian HMMs for clinical diagnoses more recently in Wang et al. (2023) and Lu et al. (2023).

While the approach we present can be considered a transparent learning algorithm, many black-box ML or artificial intelligence (AI) approaches are proving to be quite useful (e.g., Hornbrook et al. 2017, Bedoya et al. 2020, Kwon et al. 2020, Pannu et al. 2020, McLouth et al. 2021, Itzhak et al. 2023, Jha & Feng 2023). Our developed RSM extends what exists in the statistical/ML literature to build a tool that more closely addresses features of the data that are most relevant towards adequately defining and detecting patients at high risk for shock and internal bleeding. For example, the mean structure of our response model is an approximation to how an anesthesiologist would characterize a shock event (see Section 3.1). Moreover, the Bayesian framework provides an interpretable approach to quantifying the likelihood of a bleeding event based on four response outcomes (heart rate, mean arterial pressure (MAP), hemoglobin, and lactate) by way of a discrete posterior distribution of a latent state sequence across time. At any point on a discretized grid of time, the model can provide a probabilistic notion of the chance of bleeding from which a clinician can then interpret and act accordingly.

Lastly, the quality and quantity of the electronic health record data that we have gathered and curated for training and testing leads to a case study that offers real clinical insight and ramifications. Not only do these data contain vital sign and lab measurements for 33,924 patients, but they also contain detailed medication records (more in Section 2.2).

The data were retrospectively gathered and curated specifically for our study.

The remainder of the paper is structured as follows. Section 2 provides background on the medical importance of our proposed procedure (Section 2.1), a detailed description of the data (Section 2.2), and information on state-space models more broadly (Section 2.3). Section 3 provides the explicit model construction as well as the novel state-sampling algorithm (Section 3.4). Then, Section 4 presents a thorough simulation study to evaluate model performance with respect to determining/calibrating the likelihood of a bleeding event. This is followed by real data results and a case study in Section 5. Lastly, Section 6 discusses the clinical ramifications of this work as well as areas for future work.

2 Background

2.1 Medical Importance

A key problem that ICUs face is that numerous patients suffer major health complications due to bleeding and shock events that go *undetected* for too long. One study found that among patients admitted with severe trauma, one in three patients saw an intervention-time three or more hours after hospital admission, and one in six had an intervention-time six or more hours post-admission (Tran et al. 2020). The improvement of outcomes during acute bleeding and other shock states is primarily dependent on prompt diagnosis and management (Strehlow 2010), requiring "time sensitivity and patient specificity" (Convertino & Cardin 2022). In the setting of suspected hemorrhage related to trauma, diagnosis by emergency medical providers includes a Focused Assessment with Sonography in Trauma (FAST) exam and other imaging such as computed tomography (CT) (Latif et al. 2023). Resuscitation and procedural intervention, including transfusion and surgery, are based on

these findings as well as vital signs and lab work (Hooper & Armstrong 2022). In the world of trauma medicine, the "golden hour" concept, in which prioritization is given to rapid treatment with key interventions in the first hours of patient deterioration, is considered the gold standard in the management of trauma patients (Sampalis et al. 1993, 1999, Clarke et al. 2002). Similarly, time critical management of patients with sepsis according to the Surviving Sepsis guidelines has led to improved outcomes for this population (Evans et al. 2021). In patients with shock secondary to infection (i.e., sepsis), each 1-hour delay in antibiotic initiation is associated with a 10% increase in mortality (Peltan et al. 2019). In times of acute bleeding, recognition and treatment even minutes earlier may lead to substantial improvements in patient outcomes, as death from exsanguination can occur in as little as five minutes (Kotwal et al. 2018).

2.2 Data Description

The focal point of this analysis is on the 33,924 real patient encounters from the ICUs of Mayo Clinic. This dataset consists of 33 distinct types of lab measurements, vital sign recordings, and other medical descriptors of the patients over their encounters in the ICU. The data are structured in a panel-observed format with measurements discretized to a 15 minute grid. Missing data for each patient varies depending on the specific response measurement; for example, lab measurements have a high degree of missingness, whereas vital sign recordings have little to none.

Additionally, a detailed medication history is provided for each individual. Medication information is critical for detecting internal bleeding. In particular, medications affect the trends in heart rate and MAP (not hemoglobin and lactate), and depending on the dose and/or frequency with which these drugs are administered, their effects on the trends of

these vitals can be exacerbated. Therefore, by *not* accounting for medications in the model, we risk confounding trends in physiological condition based on heart rate and MAP with the possible influence of medication administration.

Further details about the data cleaning process and medication information can be found in Supplementary Materials Section 1.

2.3 Review of State-Space Models

When implementing state-space models, there usually exists two inferential interests: (1) model parameter estimates, and (2) the "most-likely" latent state sequence. For our purposes, the latter is of greater importance because learning the individual state sequences translates to learning the onset of bleeding or shock events. That said, the inference from the model parameters makes it possible to provide an interpretable probabilistic assessment of the likelihood of each latent state. This type of transparent learning algorithm is in contrast to the current status quo of black-box ML and AI approaches.

Although the HMM is the most ubiquitous form of a state-space model, its dependence structure does not adequately capture the nuanced relationship between our biological response and latent physiological states. Instead of assuming the responses are conditionally independent given the latent states (as in an HMM), we assume the responses follow an autoregressive process of order one. This describes an autoregressive HMM (AR-HMM), and applications of AR-HMMs include Ailliot & Monbet (2012), Stanculescu et al. (2013), Williams et al. (2024), among others. In addition to adding an autoregressive component to the response model, we also assume that the response at a given time instance is dependent on *all* latent states up to that given time point (explicit justification in Section 3). Therefore, rather than characterizing our model as an HMM or AR-HMM, we broadly refer

to it as an RSM. There exist many examples of RSMs, including Markov switching processes, switching autoregressive processes, and switching dynamic linear systems, among others (see Puerto-Santana et al. (2021) for an overview of the various types of RSMs). Generally, RSMs offer more model flexibility by weakening assumptions common to HMMs. Figure 1 presents a schematic of how the dependence structure differs between an HMM, an AR-HMM of order one, and our RSM.

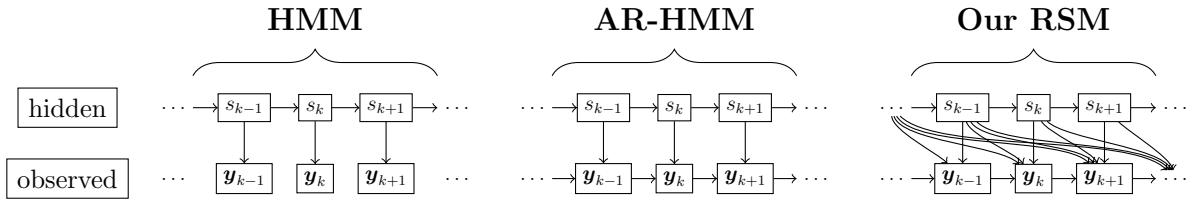

Figure 1: Schematic of the model dependence structure for an HMM, an AR-HMM of order one, and the RSM used in our approach, respectively, from left to right. Let $\boldsymbol{y}_k$ be some observed response vector at a time instance k and s_k be the corresponding latent state.

Because of the additional dependencies between the response and hidden state process, many of the well-founded inferential strategies for HMMs and AR-HMMs breakdown or are too computationally burdensome. In particular, many inferential approaches for HMMs, or AR-HMMs alike, rely on the fact that the observed process at a given time point is *only* dependent on the latent state process at that same time instant. This is not a reasonable modeling assumption in our case; hence, Section 3 precisely details the novel Markov chain Monte Carlo (MCMC) sampling routine we develop to simultaneously learn the posterior distributions of both the model parameters and the state sequences.

3 Methodology

3.1 Clinical Influence on Model Design

The manner with which we construct the model to detect internal bleeding should mimic that of a clinician's own diagnosing procedure; as such, it is necessary to understand the

indicators of possible hemorrhagic events. First, it is well understood that during a bleeding event, we expect hemoglobin to decrease, heart rate to increase, MAP to decrease, and lactate to increase. These trends serve as the canonical approach to diagnosing hemorrhage. Henceforth, our model needs to similarly track these trends, and we can do so by defining the mean of the response model as dependent on *all* latent physiological states up to a given instance of time. Second, many patients are administered drugs to stabilize their vitals. Incorporating these medications into the model is necessary in order to distinguish trends in the vitals due to physiological changes from those due to medication administration. Lastly, training this model has the additional complication that our data lack any labels indicating patient bleeding events. However, clinical expertise suggests that for historical data, if a patient received three or more red blood cell (RBC) transfusions in a 12-hour window, then some bleeding event almost certainly occurred at some point during the patient encounter. This information can serve as a partial-labeling scheme to facilitate semi-supervised learning. All of the aforementioned clinical insights shape the model construction in Section 3.2.

3.2 Model Construction

After careful consideration with clinicians, the latent physiological state-space for our RSM comprises five states: *stable* (state 1), *hemorrhage* (state 2), *recovery from hemorrhage* (state 3), *non-bleeding event* (*NBE*; state 4), and *non-bleeding event recovery* (*NBER*; state 5). State 1 describes the health condition of a patient with a low risk for shock or any complications from bleeding. The purpose of states 4 and 5 is to provide the RSM enough flexibility to characterize physiological conditions that are not state 1, nor are adequately described by states 2 or 3.

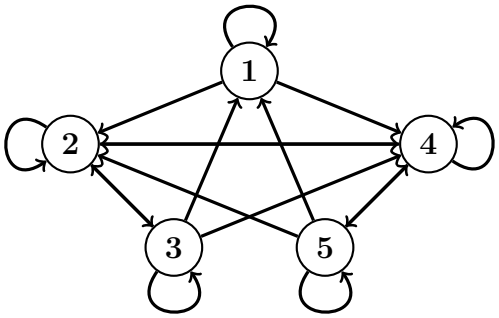

Figure 2: All allowable transitions for the five physiological states.

Let N be the number of patients, and let n_i be the number of time points observed for the i^{th} patient, where $i \in \{1, \ldots, N\}$. Let $\boldsymbol{b}_k^{(i)}$ denote the discrete random variable corresponding to the physiological state of subject i at time k, for $k \in \{1, \ldots, n_i\}$, where $\boldsymbol{b}_k^{(i)} \in \{1, \ldots, 5\}$. Let $\boldsymbol{Y}^{(i)}$ be a $4 \times n_i$ matrix with rows corresponding to the longitudinal emission variable measurements of hemoglobin, heart rate, MAP, and lactate, respectively, and columns, $\boldsymbol{y}_k^{(i)}$, corresponding to the measurements at time k.

3.2.1 Latent State Model

Since the five physiological states are non-observable, the state sequence $\boldsymbol{b}^{(i)}$ is most naturally treated as latent. Furthermore, because our data are observed every 15 minutes (i.e., on a grid), the state sequences are modeled as a discrete-time, discrete-state Markov process, with allowable transitions defined by Figure 2. The transition probability matrix, $\mathbf{P}$, characterizing the transitions between physiological states is defined as

$$\mathbf{P} := \begin{pmatrix} \frac{1}{1+e^{q_1}+e^{q_2}} & \frac{e^{q_1}}{1+e^{q_1}+e^{q_2}} & 0 & \frac{e^{q_2}}{1+e^{q_1}+e^{q_2}} & 0 \\ 0 & \frac{1}{1+e^{q_3}+e^{q_4}} & \frac{e^{q_3}}{1+e^{q_3}+e^{q_4}} & \frac{e^{q_4}}{1+e^{q_3}+e^{q_4}} & 0 \\ \frac{e^{q_5}}{1+e^{q_5}+e^{q_6}+e^{q_7}} & \frac{e^{q_6}}{1+e^{q_5}+e^{q_6}+e^{q_7}} & \frac{1}{1+e^{q_5}+e^{q_6}+e^{q_7}} & \frac{e^{q_7}}{1+e^{q_5}+e^{q_6}+e^{q_7}} & 0 \\ 0 & \frac{e^{q_8}}{1+e^{q_8}+e^{q_9}} & 0 & \frac{1}{1+e^{q_8}+e^{q_9}} & \frac{e^{q_9}}{1+e^{q_8}+e^{q_9}} \\ \frac{e^{q_{10}}}{1+e^{q_{10}}+e^{q_{11}}+e^{q_{12}}} & \frac{e^{q_{11}}}{1+e^{q_{10}}+e^{q_{11}}+e^{q_{12}}} & 0 & \frac{e^{q_{12}}}{1+e^{q_{10}}+e^{q_{11}}+e^{q_{12}}} & \frac{1}{1+e^{q_{10}}+e^{q_{11}}+e^{q_{12}}} \end{pmatrix}, \quad (1)$$

where $q_1, q_2, \ldots, q_{12}$ are linear functions of the form $q_j = \zeta_{0,j} + \zeta_{1,j} \cdot z_k$, for $j \in \{1, 2, \ldots, 12\}$, and z_k is the number of RBCs *ordered* at time point k. Note that the r^{th} row and s^{th} column of $\mathbf{P}$, denoted $\mathbf{P}_{r,s}$, is defined as $\mathbf{P}_{r,s} := P(\boldsymbol{b}_k^{(i)} = s \mid \boldsymbol{b}_{k-1}^{(i)} = r)$ for $k \in \{2, 3, \ldots, n_i\}$ and

$r, s \in \{1, 2, \ldots, 5\}$. We distinguish the number of RBCs *ordered* versus the number of RBCs *administered* because in most cases, there can exist a 30 minute delay from when RBCs are ordered versus administered. Therefore, RBC *order* times serve as the covariate in the state transition probability matrix because they mostly coincide with detectable time instances of physiological change. The prior mass function for the latent state sequence of subject i is then given by

$$p\big(\boldsymbol{b}_1^{(i)} = s_{i,1}, \ldots, \boldsymbol{b}_{n_i}^{(i)} = s_{i,n_i} \mid \boldsymbol{\pi}, \{\zeta_{0,j}, \zeta_{1,j}\}_{j=1}^{12}\big) = \boldsymbol{\pi}_{s_{i,1}} \cdot \prod_{k=2}^{n_i} \mathbf{P}_{s_{i,k-1}, s_{i,k}},$$

where $\boldsymbol{\pi} := (\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_5)^\top$ is a 5×1 vector corresponding to the discrete, initial state distribution of the Markov process, and $s_{i,1}, \ldots, s_{i,n_i} \in \{1, \ldots, 5\}$.

3.2.2 Conditional Response Model

The response, $\boldsymbol{Y}^{(i)}$, conditional on the latent states, $\boldsymbol{b}^{(i)}$, is modeled as:

$$\begin{aligned}
\boldsymbol{y}_1^{(i)} \mid \boldsymbol{b}_1^{(i)} = s_{i,1}, \boldsymbol{\alpha}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{R}, \mathbf{A}_{s_{i,1}} &\sim \mathrm{N}_4\Big(\boldsymbol{\nu}_1^{(i)},\ \boldsymbol{\Gamma}_{s_{i,1}}\Big) \\
\boldsymbol{y}_k^{(i)} \mid \boldsymbol{y}_{k-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{k}, \boldsymbol{\alpha}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{R}, \mathbf{A}_{s_{i,k}} &\sim \mathrm{N}_4\Big(\boldsymbol{\nu}_k^{(i)} + \mathbf{A}_{s_{i,k}} \cdot (\boldsymbol{y}_{k-1}^{(i)} - \boldsymbol{\nu}_{k-1}^{(i)}),\ \boldsymbol{R}\Big),
\end{aligned} \tag{2}$$

for $k \in \{2, 3, \ldots, n_i\}$, where

$$\begin{aligned}
\boldsymbol{\nu}_1^{(i)} &= g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)}) + \boldsymbol{D}_{\omega,1}^{(i)} \boldsymbol{\omega} + \boldsymbol{X}_1^{(i)} \boldsymbol{\beta}, \\
\boldsymbol{\nu}_k^{(i)} &= g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)}) + \left[\sum_{j=2}^{k} \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 2\}\right] \boldsymbol{\alpha}_{\cdot,2}^{(i)} + \cdots + \left[\sum_{j=2}^{k} \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 5\}\right] \boldsymbol{\alpha}_{\cdot,5}^{(i)} + \boldsymbol{D}_{\omega,k}^{(i)} \boldsymbol{\omega} + \boldsymbol{X}_k^{(i)} \boldsymbol{\beta}, \\
\boldsymbol{D}_{\omega,k}^{(i)} &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ d_{h,k,1}^{(i)} \cdots d_{h,k,n_{hr}}^{(i)} & \mathbf{0} \\ \mathbf{0} & d_{m,k,1}^{(i)} \cdots d_{m,k,n_{map}}^{(i)} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{X}_k^{(i)} = x_k^{(i)} \cdot \boldsymbol{I}_4.
\end{aligned} \tag{3}$$

Let $\boldsymbol{\alpha}^{(i)}$ be defined as a 4×5 matrix of random effect coefficients where the rows correspond to the four responses. The first column $\boldsymbol{\alpha}_{\cdot,1}^{(i)}$ corresponds to the subject-specific mean response when stable (i.e., state 1), and columns $\boldsymbol{\alpha}_{\cdot,2}^{(i)}, \boldsymbol{\alpha}_{\cdot,3}^{(i)}, \boldsymbol{\alpha}_{\cdot,4}^{(i)}$, and $\boldsymbol{\alpha}_{\cdot,5}^{(i)}$ correspond to

the expected change in response during each of the four non-stable states (i.e., states 2, 3, 4, and 5, respectively). The intuition for not having a slope coefficient for state 1 is because all responses should exhibit no trend when stable. The remaining terms of (2) and (3) are defined as follows: $\boldsymbol{\omega}$ is a $(n_{hr} + n_{map}) \times 1$ vector of medication effects (the first $n_{hr} = 34$ are medications affecting heart rate, and the last $n_{map} = 50$ are medications affecting MAP); $d^{(i)}_{h,k,\cdot}$ and $d^{(i)}_{m,k,\cdot}$ are the subject-specific doses for medications affecting heart rate and MAP, respectively, at time k; $\boldsymbol{\beta}$ is a 4×1 vector of coefficients defining the effects of the *administered* RBC transfusions on the mean process; $x^{(i)}_k$ is the number of RBC transfusions *administered* up to time k; $\mathbf{A}_1, \ldots, \mathbf{A}_5$ are 4×4 matrices of state-specific autocorrelation coefficients (similar to the model definition in Li-wei et al. (2015)); $\boldsymbol{R}$ is the error covariance matrix; and $g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}^{(i)}_1)$ is a random effect intercept term accounting for physiological variation upon ICU admission (see Section 3.2.4). Note that $\boldsymbol{D}^{(i)}_{\omega,k}$ has a sparse structure because medications do *not* affect hemoglobin or lactate. Lastly, as suggested in Figure 1, the mean for subject i at time k is dependent on $\{\boldsymbol{b}^{(i)}_j = s_{i,j}\}^k_{j=1}$.

Next, we assume a stable AR response by defining $\mathbf{A}_j := \text{diag}\{a_{1,j}, a_{2,j}, a_{3,j}, a_{4,j}\}$ with $a_{1,j}, a_{2,j}, a_{3,j}, a_{4,j} \in [0, 1]$ for $j \in \{1, \ldots, 5\}$. A standard assumption/property for the unconditional covariance, $\boldsymbol{\Gamma}_{s_{i,1}}$, is that it satisfies $\boldsymbol{\Gamma}_{s_{i,1}} = \mathbf{A}_{s_{i,1}} \boldsymbol{\Gamma}_{s_{i,1}} \mathbf{A}^T_{s_{i,1}} + \boldsymbol{R}$. Moreover, given the defined structure of $\mathbf{A}_{s_{i,1}}$, the j^{th} row and l^{th} column of $\boldsymbol{\Gamma}_{s_{i,1}}$ has the following form $\left[\boldsymbol{\Gamma}_{s_{i,1}}\right]_{j,l} = [\boldsymbol{R}]_{j,l}/(1 - a_{j,s_{i,1}} a_{l,s_{i,1}})$. The autocorrelation coefficient matrices are state-dependent because it is often the case that heart rhythm changes as physiological conditions change (Latif et al. 2023). Consequently, we account for heart rhythm variation by via state-dependent autocorrelation.

3.2.3 Joint Conditional Density

The joint conditional density for the data is given by

$$f\Big(\{\boldsymbol{Y}^{(i)}\}_{i=1}^{N} \mid \{\boldsymbol{b}^{(i)}\}_{i=1}^{N}, \{\boldsymbol{\alpha}^{(i)}\}_{i=1}^{N}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_5, \boldsymbol{R}\Big) = \prod_{i=1}^{N} f\Big(\boldsymbol{y}_1^{(i)} \mid \boldsymbol{b}_1^{(i)} = s_{i,1}, \boldsymbol{\alpha}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_{s_{i,1}}, \boldsymbol{R}\Big)$$
$$\times \prod_{k=2}^{n_i} f\Big(\boldsymbol{y}_k^{(i)} \mid \boldsymbol{y}_{k-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{k}, \boldsymbol{\alpha}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_{s_{i,k}}, \boldsymbol{R}\Big).$$

It is important to highlight that the parametric definition of the density is conditional on the unobserved states, $\{\boldsymbol{b}^{(i)}\}_{i=1}^{N}$. Unlike traditional HMMs, the definition of our RSM makes it infeasible to marginalize/integrate over the latent state sequences. If it were feasible to integrate out the state sequences, then standard maximum-likelihood approaches exist for efficiently learning the model parameters of HMMs, and models alike. To handle our extra model complexity, we use a Bayesian computational approach, as outlined in Section 3.3, to learn both the posterior distribution of the response model parameters, as well as the discrete posterior distribution of the latent states.

3.2.4 Accounting for Physiological Changes Before ICU Admission

It is unlikely that every patient admitted into the ICU will arrive in a stable state. More often, a patient has experienced some sort of trauma or stress prior to being admitted into the ICU. In order to account for this phenomenon, for a given subject i, the mean of the conditional response model in (2) is a function of $g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)})$, which is defined as

$$g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)}) := \boldsymbol{\alpha}_{\cdot,1}^{(i)} + t_2^{(i)} \cdot \boldsymbol{\alpha}_{\cdot,2}^{(i)} + \cdots + t_5^{(i)} \cdot \boldsymbol{\alpha}_{\cdot,5}^{(i)}, \tag{4}$$

where $t_2^{(i)}, t_3^{(i)}, t_4^{(i)}$, and $t_5^{(i)}$ represent the time spent in states 2, 3, 4, and 5, respectively, up to the initial ICU observation. Because no data exists before a subject's ICU admission, estimating $g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)})$ is not feasible. However, not accounting for this added variability in the initial observation can lead to challenges in the estimation of other model parameters.

One could naively assume either $t_2^{(i)} = t_3^{(i)} = t_4^{(i)} = t_5^{(i)} = 0$, $\forall i$, or

$$g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)}) := \boldsymbol{\alpha}_{\cdot,1}^{(i)} + \mathbf{1}\{\boldsymbol{b}_1^{(i)} = 2\} \cdot \boldsymbol{\alpha}_{\cdot,2}^{(i)} + \cdots + \mathbf{1}\{\boldsymbol{b}_1^{(i)} = 5\} \cdot \boldsymbol{\alpha}_{\cdot,5}^{(i)}.$$

Both of these simplifications, however, would conflate the variability from the random effect, $\boldsymbol{\alpha}^{(i)}$, with the variability due to non-zero $t_2^{(i)}, \ldots, t_5^{(i)}$.

We can instead *approximate* the conditional density for the data in the following manner. Given the initial observation, $\boldsymbol{y}_1^{(i)}$, and model parameters, $\boldsymbol{\omega}$ and $\boldsymbol{\beta}$, we can approximate $g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)}) \approx \boldsymbol{y}_1^{(i)} - \boldsymbol{D}_{\omega,1}^{(i)} \cdot \boldsymbol{\omega} - \boldsymbol{X}_1^{(i)} \cdot \boldsymbol{\beta}$. Hence, define the approximate conditional model:

$$\begin{aligned}
\boldsymbol{\gamma}^{(i)} \mid \boldsymbol{y}_1^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{G} &\sim \mathrm{N}_4(\boldsymbol{y}_1^{(i)} - \boldsymbol{D}_{\omega,1}^{(i)} \cdot \boldsymbol{\omega} - \boldsymbol{X}_1^{(i)} \cdot \boldsymbol{\beta},\ \boldsymbol{G}) \\
\boldsymbol{y}_1^{(i)} \mid \boldsymbol{\gamma}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{R}, \mathbf{A}_{s_{i,1}} &\sim \mathrm{N}_4\Big(\boldsymbol{\nu}_1^{(i)}, \boldsymbol{\Gamma}_{s_{i,1}}\Big) \qquad (5) \\
\boldsymbol{y}_k^{(i)} \mid \boldsymbol{y}_{k-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=2}^k, \boldsymbol{\alpha}_*^{(i)}, \boldsymbol{\gamma}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{R}, \mathbf{A}_{s_{i,k}} &\sim \mathrm{N}_4\Big(\boldsymbol{\nu}_k^{(i)} + \mathbf{A}_{s_{i,k}} \cdot (\boldsymbol{y}_{k-1}^{(i)} - \boldsymbol{\nu}_{k-1}^{(i)}), \boldsymbol{R}\Big),
\end{aligned}$$

for $k \in \{2, 3, \ldots, n_i\}$, where $\boldsymbol{G}$ is a new covariance matrix to learn,

$$\begin{aligned}
\boldsymbol{\nu}_1^{(i)} &= \boldsymbol{\gamma}^{(i)} + \boldsymbol{D}_{\omega,1}^{(i)} \cdot \boldsymbol{\omega} + \boldsymbol{X}_1^{(i)} \cdot \boldsymbol{\beta}, \\
\boldsymbol{\nu}_k^{(i)} &= \boldsymbol{\gamma}^{(i)} + \boldsymbol{D}_{\alpha,k}^{(i)} \cdot \mathrm{vec}\Big(\boldsymbol{\alpha}_*^{(i)}\Big) + \boldsymbol{D}_{\omega,k}^{(i)} \cdot \boldsymbol{\omega} + \boldsymbol{X}_k^{(i)} \cdot \boldsymbol{\beta}, \\
\boldsymbol{D}_{\alpha,k}^{(i)} &= \boldsymbol{I}_4 \otimes \left(\sum_{j=2}^k \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 2\} \quad \sum_{j=2}^k \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 3\} \quad \sum_{j=2}^k \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 4\} \quad \sum_{j=2}^k \mathbf{1}\{\boldsymbol{b}_j^{(i)} = 5\} \right),
\end{aligned}$$

$\mathrm{vec}(\cdot)$ denotes the operation of forming a column vector from a matrix by stacking the columns of the matrix from left to right, and $\boldsymbol{\alpha}_*^{(i)}$ is equal to $\boldsymbol{\alpha}^{(i)}$ with the first column removed and then transposed (i.e., the four rows of $\boldsymbol{\alpha}_*^{(i)}$ correspond to the four non-stable states and the four columns correspond to the four responses). Based on Bayesian computation routines, we can sample $\boldsymbol{\gamma}^{(i)}$ to then compute the conditional density.

To further justify this approach, Supplementary Materials Section 9 offers an additional, simpler simulation study. The purpose of this simulation is to illustrate that although (5) is an *approximate* conditional density, it is still effective at learning parameter estimates and

recovering the underlying state sequences. This additional simulation also showcases the inferential consequences of *not* accounting for pre-ICU-admission physiological changes.

3.3 Bayesian Computation

With the updated response model in (5), the approximate joint posterior distribution is

$$\begin{aligned}
&\pi\Big(\{\boldsymbol{b}^{(i)}\}_{i=1}^{N}, \{\boldsymbol{\alpha}_*^{(i)}\}_{i=1}^{N}, \{\boldsymbol{\gamma}^{(i)}\}_{i=1}^{N}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_5, \boldsymbol{R}, \boldsymbol{G}, \widetilde{\boldsymbol{\alpha}}_*, \boldsymbol{\Upsilon}_\alpha, \boldsymbol{\zeta}, \boldsymbol{\pi} \;\Big|\; \{\boldsymbol{Y}^{(i)}\}_{i=1}^{N}\Big) \\
&\quad \propto \Big\{ \prod_{i=1}^{N} \pi\Big(\boldsymbol{\alpha}_*^{(i)} \mid \widetilde{\boldsymbol{\alpha}}_*, \boldsymbol{\Upsilon}_\alpha\Big) \cdot \pi\Big(\boldsymbol{\gamma}^{(i)} \mid \boldsymbol{y}_1^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{G}\Big) \cdot \boldsymbol{\pi}_{s_{i,1}} \cdot f\Big(\boldsymbol{y}_1^{(i)} \mid \boldsymbol{\gamma}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_{s_{i,1}}, \boldsymbol{R}\Big) \\
&\qquad \times \prod_{k=2}^{n_i} \mathbf{P}_{s_{i,k-1}, s_{i,k}} \cdot f\Big(\boldsymbol{y}_k^{(i)} \mid \boldsymbol{y}_{k-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{k}, \boldsymbol{\alpha}_*^{(i)}, \boldsymbol{\gamma}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_{s_{i,k}}, \boldsymbol{R}\Big) \Big\} \\
&\qquad \times \pi(\boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_5, \boldsymbol{R}, \boldsymbol{G}, \widetilde{\boldsymbol{\alpha}}_*, \boldsymbol{\Upsilon}_\alpha, \boldsymbol{\zeta}),
\end{aligned}$$

where the random effect and selected prior distributions are

$$\begin{aligned}
&\mathrm{vec}(\boldsymbol{\alpha}_*^{(i)}) \sim \mathrm{N}_{16}(\mathrm{vec}(\tilde{\boldsymbol{\alpha}}_*),\ \boldsymbol{\Upsilon}_\alpha), && \mathrm{vec}(\tilde{\boldsymbol{\alpha}}_*) \sim \mathrm{N}_{16}(\mathrm{vec}(\tilde{\boldsymbol{\alpha}}_0),\ \boldsymbol{\Sigma}_\alpha), \\
&\boldsymbol{\omega} \sim \mathrm{N}_{n_{hr}+n_{map}}(\boldsymbol{\omega}_0,\ \boldsymbol{\Sigma}_\omega), && \boldsymbol{\Upsilon}_\alpha \sim \mathrm{InvWish}(\boldsymbol{\Psi}_\alpha,\ \nu_\alpha), \\
&\boldsymbol{\beta} \sim \mathrm{N}_4(\boldsymbol{\beta}_0,\ \boldsymbol{\Sigma}_\beta), && \boldsymbol{G} \sim \mathrm{InvWish}(\boldsymbol{\Psi}_G,\ \nu_G), \\
&\boldsymbol{R} \sim \mathrm{InvWish}(\boldsymbol{\Psi}_R,\ \nu_R).
\end{aligned}$$

Gaussian/multivariate Gaussian priors are placed on the parameters or transformations of the parameters that are not displayed here. Parameter estimation is done using a Metropolis-within-Gibbs MCMC sampling algorithm; derivations and details are provided in Supplementary Materials Section 2.

Precise prior specifications are found in Supplementary Materials Section 3, yet the following is a justification for some of the stronger priors. Notably, a diffuse prior on the noise parameter $\boldsymbol{R}$ leads to a level of noise that inhibits the MCMC algorithm from accepting any states other than state 1. The intuition is that for a sufficiently large error

variance, any change in the vitals is attributed to random chance rather than a physiological change. Consequently, we set the Inverse-Wishart prior degrees of freedom for $\boldsymbol{R}$ to $\nu_R = 2 \cdot \sum_{i=1}^{N} n_i$ and the prior scale matrix to $\boldsymbol{\Psi}_R = \nu_R \cdot \text{diag}\{\frac{1}{2}, \frac{3}{2}, \frac{3}{2}, \frac{1}{2}\}$. Next, $\boldsymbol{\omega}$ quantifies the medication effects; in particular, these medications are characterized as "uppers" or "downers" depending on their effects on heart rate and MAP. Hence, the specified prior on $\boldsymbol{\omega}$ encourages the medication effects for "uppers" to be positive and "downers" to be negative. The medication effects for Norepinephrine and Dexmedetomidine, in particular, necessitate stronger priors to estimate the correct sign.

Recall from Section 3.2.4 that the approximate conditional density is a function of the latent states and the random effects. Computing this approximate density is feasible because at every iteration of the MCMC, we sample from the conditional posterior distributions of $\{\boldsymbol{b}^{(i)}\}_{i=1}^{N}$, $\{\boldsymbol{\alpha}_*^{(i)}\}_{i=1}^{N}$, and $\{\boldsymbol{\gamma}^{(i)}\}_{i=1}^{N}$, respectively. For the purposes of sampling the random effect $\{\boldsymbol{\alpha}_*^{(i)}\}_{i=1}^{N}$, we implement a rejection sampling strategy. In particular, to better classify the canonical characteristics of states 2 and 3, we sample each $\boldsymbol{\alpha}_*^{(i)}$ according to its Gibbs update, but only "accept" if the proposed slopes for states 2 and 3 coincide with the expected hemodynamics with respect to the values being positive or negative. Lastly, for the Metropolis-Hastings (MH) updates, an adaptive proposal strategy is implemented during the burnin phase to improve the mixing/efficiency of the sampling by learning the proposal correlation structures and tuning the proposal variances/scales.

3.4 State-Sampling Algorithm

The use of Bayesian estimation methods in state-space models to learn the posterior distribution of the latent state sequences has been studied in the literature before (e.g., Carter & Kohn 1994, Djuric & Chun 2002, Scott 2002, Turek et al. 2016, Triantafyllopoulos et al.

2021, Williams et al. 2024). A detailed review of the various MCMC sampling approaches for the latent state process of an HMM is found in Fearnhead & Sherlock (2025). In our case, however (i.e., beyond the scope of HMMs or AR-HMMs), we have the additional constraint that the response model for a given time, k, is dependent on *all* latent states up to time k. As such, the complex dependence structure coupled with the high-dimensional parameter space necessitates a novel and clever state-sampling procedure beyond a naive Gibbs or MH update; otherwise, learning the discrete posterior distribution of the latent state sequences quickly becomes computationally prohibitive.

3.4.1 Existing State-Sampling Methods

Let $p(\boldsymbol{b}_k^{(i)} = s_{i,k}, \ldots, \boldsymbol{b}_{k+p-1}^{(i)} = s_{i,k+p-1} \mid \boldsymbol{b}_{k-1}^{(i)} = s_{i,k-1}, \boldsymbol{b}_{k+p}^{(i)} = s_{i,k+p}, \ \ldots)$ denote the full conditional distribution of the latent state sequence $\boldsymbol{b}_k^{(i)}, \ldots, \boldsymbol{b}_{k+p-1}^{(i)}$ for subject i at starting time k, where $p \in \mathbb{Z}^+$. This is a discrete probability mass function where the realizations of $\boldsymbol{b}_k^{(i)}, \ldots, \boldsymbol{b}_{k+p-1}^{(i)}$ are dependent on the values of $\boldsymbol{b}_{k-1}^{(i)}$ and $\boldsymbol{b}_{k+p}^{(i)}$. With the restrictions on the allowable state transitions, let $\mathcal{B}_{s_{i,k-1},s_{i,k+p}}^{(p)}$ denote the set of possible state p-tuples for $\boldsymbol{b}_k^{(i)}, \ldots, \boldsymbol{b}_{k+p-1}^{(i)}$ given $\boldsymbol{b}_{k-1}^{(i)} = s_{i,k-1}$ and $\boldsymbol{b}_{k+p}^{(i)} = s_{i,k+p}$ (e.g., $\mathcal{B}_{1,3}^{(1)} = \{\{2\}\}$ and $\mathcal{B}_{1,3}^{(2)} = \{\{1,2\},\{2,2\},\{2,3\},\{4,2\}\}$). In the special edge cases of $k = 1$ or $k = n_i - p + 1$, let $\mathcal{B}_{\cdot,s_{i,k+p}}^{(p)}$ and $\mathcal{B}_{s_{i,k-1},\cdot}^{(p)}$ define the set of possible state p-tuples, respectively.

Next, the state proposal distribution is derived by assigning probabilities to each element of $\mathcal{B}_{s_{i,k-1},s_{i,k+p}}^{(p)}$. There exists a variety of approaches for defining the state proposal distribution in the literature; we will focus on two examples. The simplest approach is to assign equal probability to each element of $\mathcal{B}_{s_{i,k-1},s_{i,k+p}}^{(p)}$ leading to an MH sampling procedure. This approach is seen in Williams et al. (2024) and will be termed *approach (A)*. Alternatively, in a Gibbs update, the proposal distribution can incorporate information from the data as well as the current response parameter values in the MCMC iteration.

This state-sampling approach is seen in Albert & Chib (1993), Robert et al. (1993), Robert & Titterington (1998), and will be termed *approach (B)*. Their precise details are provided in Supplementary Materials Section 7.

While approaches (A) and (B) are mathematically different proposal strategies, they both require computing $\mathcal{B}^{(p)}_{s_{i,k-1},s_{i,k+p}}$ for any combination of $s_{i,k-1}, s_{i,k+p} \in \{1,\ldots,5\}$. Additionally, a value for p must be selected which, as we will demonstrate, can lead to computational challenges. In the case of approach (B), the Gibbs updates require looping over all elements of $\mathcal{B}^{(p)}_{s_{i,k-1},s_{i,k+p}}$. As Table 1 shows, an increase in p leads to an exponential increase in the cardinality of $\mathcal{B}^{(p)}_{s_{i,k-1},s_{i,k+p}}$ and consequently an exponential increase in compute time. Approach (A) does not exhibit the same computational inefficiency; however, it is subject to a statistical/algorithmic inefficiency in the sense that as the cardinality of $\mathcal{B}^{(p)}_{s_{i,k-1},s_{i,k+p}}$ increases, the probability of randomly selecting a "good" proposal sequence decreases because each element in the set is assigned equal probability. In other words, each iteration of approach (B) is slower, but potentially many more iterations of approach (A) may be required. A more nuanced state-sampling strategy is needed.

	$p=1$	$p=2$	$p=4$	$p=6$	$p=8$	$p=10$
$\lvert\mathcal{B}^{(p)}_{1,1}\rvert$	1	3	37	395	4221	45123
$\lvert\mathcal{B}^{(p)}_{2,3}\rvert$	2	5	48	509	5436	58105
$\lvert\mathcal{B}^{(p)}_{5,4}\rvert$	4	13	137	1465	15661	167413

Table 1: Examples of how the cardinality of $\mathcal{B}^{(p)}_{s_{i,k-1},s_{i,k+p}}$ changes as a function of p.

3.4.2 Our State-Sampling Procedure

Our novel state-sampling strategy can be found in Algorithm 1 from Supplementary Materials Section 5. To contextualize this algorithm, the state-sampling works as follows. The algorithm is an MH sampling routine. There exists flexibility in the number of time points, p, over which the state-space is sampled. In one extreme, for each subject, i, we can take

$p = n_i$ and propose full state sequences (special case Algorithm 2), or we can propose states over blocks of time instances in sequence with minimum $p = 2$ (special case Algorithm 3).

For each step of the MCMC sampling routine, we sample $p \sim \text{uniform}\{2, \ldots, 50\}$. For a given subject, i, and starting time, $k \in \{1, \ldots, n_i - 2\}$, Algorithm 1 makes state proposals in accordance to the likelihood and transition probabilities across time points $\{k, k+1, \ldots, k_{\max}\}$ where $k_{\max} = \min\{k + p - 1, n_i\}$. For most values of p and k, the discrete proposal distribution for $\{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=k}^{k_{\max}}$ is defined as

$$
\begin{aligned}
&q\Big(\{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=k}^{k_{\max}} \mid \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j \notin \{k,\ldots,k_{\max}\}}, \boldsymbol{Y}^{(i)}, \boldsymbol{\alpha}_*^{(i)}, \boldsymbol{\gamma}^{(i)}, \boldsymbol{\omega}, \boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_5, \boldsymbol{R}, \boldsymbol{\zeta}, \boldsymbol{\pi}\Big) \\
&:= \begin{cases}
\dfrac{\boldsymbol{\pi}_{s_{i,1}} \cdot f\left(\boldsymbol{y}_1^{(i)} | \boldsymbol{b}_1^{(i)} = s_{i,1},\, \text{rest}\right)}{\sum_{m=1}^{5} \boldsymbol{\pi}_m \cdot f\left(\boldsymbol{y}_1^{(i)} | \boldsymbol{b}_1^{(i)} = m,\, \text{rest}\right)} \displaystyle\prod_{t=k+1}^{k_{\max}-2} \frac{\mathbf{P}_{s_{i,t-1}, s_{i,t}} \cdot f\left(\boldsymbol{y}_t^{(i)} | \boldsymbol{y}_{t-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{t},\, \text{rest}\right)}{\sum_{m=1}^{5} \mathbf{P}_{s_{i,t-1}, m} \cdot f\left(\boldsymbol{y}_t^{(i)} | \boldsymbol{y}_{t-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{t-1}, \boldsymbol{b}_t^{(i)} = m,\, \text{rest}\right)} \\
\qquad \times \left| \mathcal{B}^{(2)}_{s_{i,k_{\max}-2}, s_{i,k_{\max}+1}} \right|^{-1} & , \; k = 1 \\
\displaystyle\prod_{t=k}^{k_{\max}-2} \frac{\mathbf{P}_{s_{i,t-1}, s_{i,t}} \cdot f\left(\boldsymbol{y}_t^{(i)} | \boldsymbol{y}_{t-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{t},\, \text{rest}\right)}{\sum_{m=1}^{5} \mathbf{P}_{s_{i,t-1}, m} \cdot f\left(\boldsymbol{y}_t^{(i)} | \boldsymbol{y}_{t-1}^{(i)}, \{\boldsymbol{b}_j^{(i)} = s_{i,j}\}_{j=1}^{t-1}, \boldsymbol{b}_t^{(i)} = m,\, \text{rest}\right)} \cdot \left| \mathcal{B}^{(2)}_{s_{i,k_{\max}-2}, s_{i,k_{\max}+1}} \right|^{-1} & , \; \text{else}
\end{cases}
\end{aligned}
\tag{6}
$$

See Supplementary Materials Section 6 for other/all cases.

Similar to approach (B), Algorithm 1 proposes state sequences in proportion to the likelihood given the model parameters of the current MCMC iteration. Whereas approach (B) has computational complexity $\mathcal{O}(5^p)$, our Algorithm 1 achieves complexity $\mathcal{O}(5(p-2) + 5^2)$. Algorithm 1 shares a similar structure to the "forward-backward Gibbs sampler" seen in Scott (2002).

To better illustrate the efficiency of Algorithm 1, Table 2 presents the median compute time for one iteration of an MCMC sampling routine using Algorithm 1 versus approaches (A) and (B) on a dataset with 10 simulated subjects. As a baseline, we initialize each algorithm with the state sequence having all states being state 1. For approach (A), the accuracy is maximized when $p = 4$, and decreases for larger p. This highlights the intuition given in Section 3.4.1 that as the cardinality of $\mathcal{B}^{(p)}_{s_{i,k-1}, s_{i,k+p}}$ increases, the true state sequence has a lower chance of being proposed via uniform sampling at each iteration, and so more iterations of the algorithm are needed to happen upon it. Next, we see that approach

(B) with $p = 8$ leads to the second highest accuracy, though, it only completed 36 out of the 100 iterations in 12 hours. When compared to Algorithm 1 [which obtained the highest accuracy], approach (B) with $p = 8$ is 36,394 times slower. After the 100 iterations, Algorithm 1 achieves the fastest median compute time *and* the highest state accuracy.

Approach	Time (sec.)	Percent correct
A ($p = 2$)	0.5598	0.6340
A ($p = 4$)	0.8022	0.6481
A ($p = 6$)	0.8813	0.6114
A ($p = 8$)	1.0659	0.5472
A ($p = 10$)	4.5732	0.5240
B ($p = 2$)	1.9612	0.6939
B ($p = 4$)	24.966	0.7673
B ($p = 6$)	267.83	0.8082
B ($p = 8$)	2653.1*	0.8392*
B ($p = 10$)	29333.9*	0.8258*
Algorithm 1	0.0729	0.8766

Table 2: Median compute times (in seconds) over 100 iterations of each state-sampling algorithm and the accuracy of each sampling algorithm after the 100 iterations. Accuracy is defined as the proportion of time instances where the posterior modal state (after 100 runs) correctly corresponds to the true state. The $(*)$ indicates that the 100 iterations did not finish within 12 hours. In particular, the Gibbs update (approach (B)) with $p = 10$ only completed *three* iterations before the algorithm timed out. Note that Algorithm 1 does not change with respect to p because at each iteration of the MCMC routine, a new p is randomly selected from $\{2, \ldots, 50\}$.

Lastly, recall from Section 3.1 that we can use the number of RBC transfusions to serve as partial labels in our training data. In particular, if a patient receives three or more RBC transfusions in any 12 hour window, they have likely suffered some sort of hemorrhagic event. This indicator is built into our state-sampling in the following manner. For patient encounters satisfying this rule, when a state sequence is proposed, the sequence must contain state 2 at some point before the last of the RBC transfusion order times in the 12 hour window, else rejected. In addition to this heuristic, a small subset of the patient encounters have been evaluated by our three anesthesiologists to provide a yes/no answer as to whether the patient suffered from internal bleeding during their encounter. If the patient was clinically annotated "yes," then proposed state sequences, again, must contain at least one state 2. Conversely, if the patient was clinically annotated "no," then

the only allowable states are 1, 4, and 5. Note, however, that these rules are removed for the case study in Section 5.2.

4 Simulation Study

The lack of state labels in the real data necessitates a thorough simulation study to test our model's ability to detect internal bleeding. This simulation study serves a crucial role as it lays the foundation for how we interpret the discrete posterior probabilities of the physiological states. If we consider applying our approach in a hospital setting, we need to quantify what a "high" posterior probability of internal bleeding is such that we can adequately alarm for clinical intervention. Therefore, this simulation will determine a threshold/cutoff for the posterior probability of internal bleeding that is calibrated to balance the sensitivity of identifying state 2 with minimizing the number of false alarms. This threshold is then applied in our real case study in Section 5.2.

4.1 Data Generating Mechanism

In order to simulate data that closely resembles the observed training data, we take the exact covariate information, length of patient encounters, and level of missingness from the real data. The true parameter values are determined from training the model on the real data as in Section 5; details are provided in Supplementary Materials Section 8. Given the covariate information and "true" parameter values, the latent state sequences and longitudinal response measurements are simulated as follows.

The data generating mechanism is identical to the model definition provided in Sections 3.2.1 and 3.2.2, with one exception. First, for each subject $i \in \{1, \ldots, N\}$, the latent state sequence is simulated according to the transition probability matrix defined in (1). Recall

from Section 3.2.4 that our model accounts for possible physiological changes prior to a patient's ICU admission. Thus, we need to simulate values for $t_2^{(i)}, \ldots, t_5^{(i)}$ from (4). With n_i representing the length of the observed patient encounter, we generate a state sequence of length $n_i + m_i$, where $m_i \sim \text{uniform}\{0, 1, \ldots, 50\}$. The first m_i time points allow for possible state changes before the start of the patient encounter, with the initial state always being state 1. The first m_i states are generated according to (1), except we assume $q_j = \zeta_{0,j}$, $\forall j \in \{1, \ldots, 12\}$. The last n_i time points correspond to the observed process. Let $\boldsymbol{b}_{long}^{(i)}$ denote the state sequence of length $n_i + m_i$, and $\boldsymbol{b}^{(i)}$ represent the true state sequence for subject i (i.e., the last n_i states of $\boldsymbol{b}_{long}^{(i)}$). Then $t_l^{(i)} := \sum_{k=1}^{m_i+1} \mathbf{1}\{\boldsymbol{b}_{long,\,k}^{(i)} = l\}$ for $l \in \{2, 3, 4, 5\}$ is used to define $g(\boldsymbol{\alpha}^{(i)}, \boldsymbol{b}_1^{(i)})$. Then, we generate the responses according to model (2), setting the proportion of missingness to that of the actual patient encounter.

One hundred datasets are simulated, with each containing 1,000 patient encounters (500 training and 500 testing), and the MCMC algorithm is run for 10,000 steps with the first 5,000 discarded as burnin.

4.2 Interpretation of Results and Posterior Probabilities

One approach to evaluate model performance in identifying the latent state sequence for each subject is to compare the posterior modal state at each time point to the true state (as previously done in Table 2). Using this metric, the median accuracy across all datasets is 0.7393. However, for our purposes, we are more interested in how precise our model is at detecting specifically state 2. Posterior probabilities are not necessarily calibrated to frequentist/aleatory probabilities, and so we must determine the minimum posterior probability of bleeding threshold $c \in [0, 1]$ that best identifies the onset of state 2.

Figure 3 presents the ROC curve for the identification of state 2 for *one* of the 100

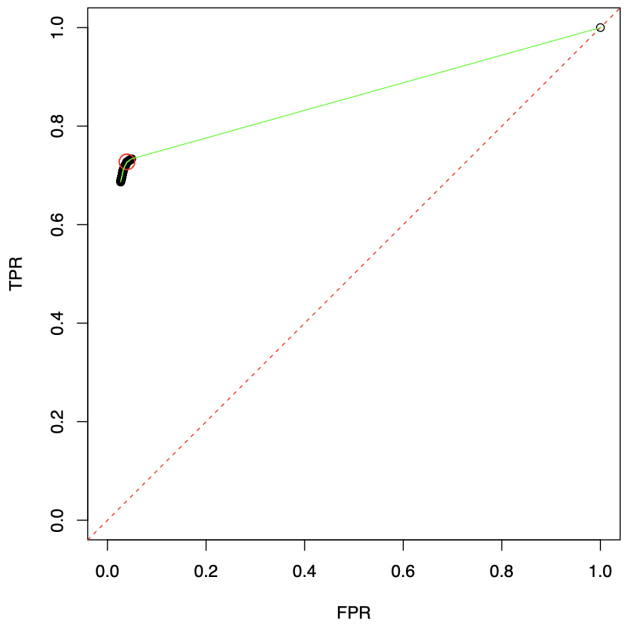

Figure 3: ROC curve depicting the sensitivity and specificity of identifying state 2. The red circle indicates the posterior probability threshold $c \in [0, 1]$ that minimizes FPR - TPR.

simulated datasets. For each simulated dataset, a value of c is determined by finding the one that minimizes the difference between the false positive rate (FPR) and true positive rate (TPR). Using the median value of c across all datasets as a threshold, we get $\hat{c} = 0.0465$. The median out-of-sample AUC (computed using the trapezoidal integral approximation) across all simulations is 0.8419. Using the threshold $\hat{c} = 0.0465$, the median sensitivity of correctly identifying state 2 is 0.7270, and the median specificity of correctly identifying *not* state 2 is 0.9621. Table 3 provides additional summary statistics. Trace plots and box plots for the simulation study are provided in Supplementary Materials Section 11.

	Q1	Median	Mean	Q3
threshold value c	0.0208	0.0465	0.0543	0.0700
out-of-sample AUC	0.8303	0.8419	0.8400	0.8501
sensitivity of state 2	0.7027	0.7270	0.7204	0.7412
specificity of not state 2	0.9581	0.9621	0.9622	0.9655

Table 3: Summary statistics across 100 simulated datasets. The sensitivity and specificity calculations are done using $\hat{c} = 0.0465$.

An important takeaway from this simulation study is how small the threshold for the

posterior probability of bleeding is. Recall the true parameter values for the simulation come from the trained model in Section 5. Based on the parameter estimates in Table 4, the probability of transitioning out of states 1 or 5 is relatively *low*, whereas the probability of transitioning out of states 2, 3, or 4 is *high*. Namely, state 2 is a rare event meaning that it is indicated by even a small posterior probability in the fitted RSM.

5 Real Data Analysis

5.1 Results from Training Data

Recall from Section 3 that $\tilde{\boldsymbol{\alpha}}_*$ and $\boldsymbol{\Upsilon}_\alpha$ are the random effect mean and variance terms, respectively, which characterize the expected change in vital response based on the current and previous latent states. In addition, recall that $\boldsymbol{\zeta}$ are the transition rate coefficients characterizing the evolution of the latent state process, and $\{\mathbf{A}_j\}_{j=1}^5$ are the state-dependent AR coefficient matrices quantifying the autocorrelation in response outcomes over time. The estimates of $\tilde{\boldsymbol{\alpha}}_*$, $\boldsymbol{\Upsilon}_\alpha$, $\boldsymbol{\zeta}$, and $\{\mathbf{A}_j\}_{j=1}^5$ offer particularly useful insight into understanding the physiological characteristics of onset and offset of bleeding. Figure 4 visualizes sampled random effect coefficients from the fitted RSM. Tables 4 and 5 provide the posterior median estimates and 95% credible intervals for $\boldsymbol{\zeta}$ and $\{\mathbf{A}_j\}_{j=1}^5$, respectively.

Figure 4 and Table 4 characterize the expected phenotypes for the different latent states. Recall that during a bleeding event (state 2), it is expected that hemoglobin decreases, heart rate increases, MAP decreases, and lactate increases. The relation to zero of the violin plots in Figure 4 are consistent with these behaviors. We must acknowledge that the rate of bleeding varies for every patient experiencing blood loss, and changes in vital signs and lab values are frequently delayed. Nonetheless, to summarize a typical presentation

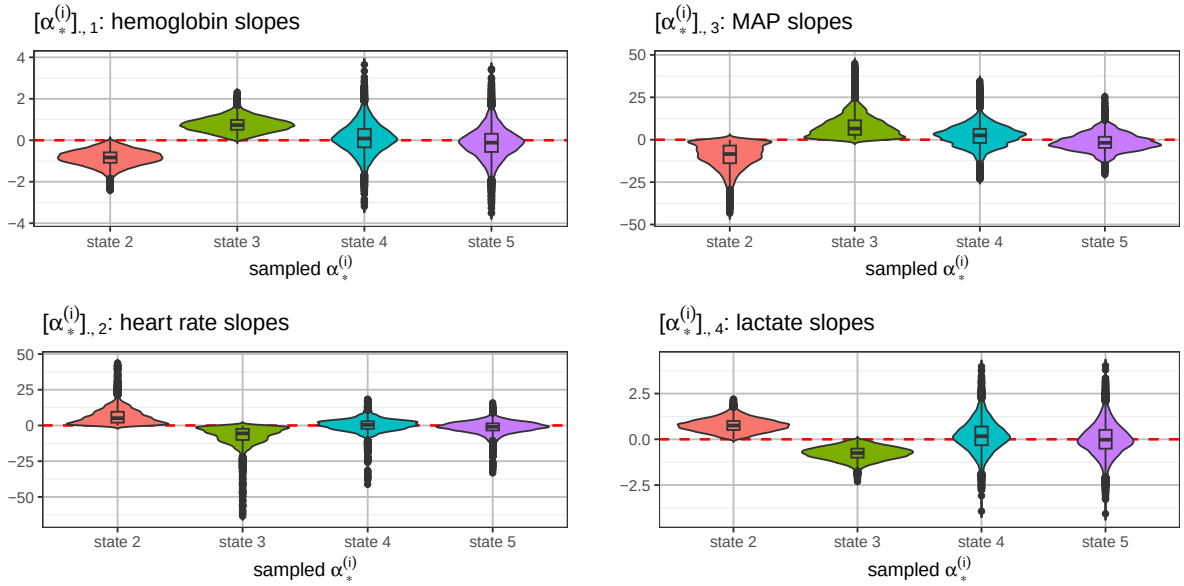

Figure 4: Sampled random effect coefficients across all 500 training subjects for the last MCMC iteration.

of a patient suffering from a clinically significant bleeding event, the fitted model suggests the following changes over a 15 minute period: a decrease in hemoglobin of 0.846 g/dL, an increase in heart rate of 6.559 bpm, a decrease in MAP of 9.623 mmHg, and an increase in lactate of 0.749 mmol/L. Figure 4 also illustrates that the inter-individual variability in hemorrhage phenotype is quite large. For example, the violin plots for the slope coefficients on heart rate and MAP during bleeding suggest that many exhibit trends similar to the estimated parent mean $\tilde{\boldsymbol{\alpha}}_*$; however, there are also some subjects exhibiting changes in heart rate and MAP of over 40 units in magnitude in a 15 minute period. Additionally, for a given vital, we see that the sampled slope coefficients for state 3 are roughly a mirrored image of the sampled slopes for state 2, which aligns with the intuition that state 3 is characterizing the recovery from bleeding.

In terms of states 4 and 5, namely NBE and NBER, we see more variation in the sampled random effect slopes in Figure 4 as compared to states 2 and 3. It is difficult to precisely characterize a physiological interpretation for states 4 and 5 because while they can account for more variable hemodynamic changes in vitals, they also appear to

State Transitions		
$\mathbf{1 \to 2}$	$\mathbf{1 \to 4}$	$\mathbf{2 \to 3}$
$\widehat{\zeta}_{0,1} = -2.996[-3.069, -2.924]$	$\widehat{\zeta}_{0,2} = -2.127[-2.174, -2.078]$	$\widehat{\zeta}_{0,3} = 1.94[1.824, 2.055]$
$\widehat{\zeta}_{1,1} = 0.542[0.172, 0.851]$	$\widehat{\zeta}_{1,2} = 0.231[-0.188, 0.55]$	$\widehat{\zeta}_{1,3} = -0.513[-1.064, 0.087]$
$\mathbf{2 \to 4}$	$\mathbf{3 \to 1}$	$\mathbf{3 \to 2}$
$\widehat{\zeta}_{0,4} = 0.616[0.48, 0.754]$	$\widehat{\zeta}_{0,5} = 1.095[0.979, 1.211]$	$\widehat{\zeta}_{0,6} = 0.145[0.016, 0.274]$
$\widehat{\zeta}_{1,4} = 0.105[-0.449, 0.701]$	$\widehat{\zeta}_{1,5} = -0.763[-1.793, 0.144]$	$\widehat{\zeta}_{1,6} = -0.136[-0.989, 0.662]$
$\mathbf{3 \to 4}$	$\mathbf{4 \to 2}$	$\mathbf{4 \to 5}$
$\widehat{\zeta}_{0,7} = 0.301[0.158, 0.438]$	$\widehat{\zeta}_{0,8} = -1.056[-1.143, -0.97]$	$\widehat{\zeta}_{0,9} = 0.726[0.674, 0.78]$
$\widehat{\zeta}_{1,7} = -0.594[-1.714, 0.348]$	$\widehat{\zeta}_{1,8} = 0.508[-0.056, 1.095]$	$\widehat{\zeta}_{1,9} = 0.054[-0.4, 0.681]$
$\mathbf{5 \to 1}$	$\mathbf{5 \to 2}$	$\mathbf{5 \to 4}$
$\widehat{\zeta}_{0,10} = -0.178[-0.232, -0.121]$	$\widehat{\zeta}_{0,11} = -1.981[-2.101, -1.871]$	$\widehat{\zeta}_{0,12} = -0.483[-0.549, -0.419]$
$\widehat{\zeta}_{1,10} = -0.449[-1.008, 0.007]$	$\widehat{\zeta}_{1,11} = 0.156[-0.367, 0.557]$	$\widehat{\zeta}_{1,12} = -0.026[-0.444, 0.318]$

Table 4: The posterior median estimates and 95% credible sets for the transition rate parameters.

be indicating more prolonged physiological conditions, with reference to the transition rate coefficient estimates in Table 4.

A basic reading of Table 4 indicates that transitioning from state 1 to either state 2 or 4 (in any 15 minute unit of time) is rarer than any other valid state transition. Furthermore, state $1 \to 2$ is *less* likely than state $1 \to 4$, i.e., consistent with state 2 characterizing only internal bleeding, whereas state 4 is meant to characterize all other physiological changes.

Lastly, Table 5 provides an additional perspective on how best to interpret the five states. Clinical intuition suggests that patients' vitals are more highly autocorrelated when stable, whereas they appear more sporadic and less autocorrelated in the presence of shock. As suggested in Table 5 as well as by our *a priori* belief, heart rate, MAP, and lactate are highly autocorrelated when a patient is stable. However, for states 2, 3, 4, and 5, the autocorrelation between successive heart rate and MAP measurements is much less. Because all states aside from state 1 represent non-stable physiological conditions, the low degree of autocorrelation aligns with our previous intuition. As for the autocorrelation coefficients on hemoglobin and lactate, we do not see as drastic a change across the five states as compared with heart rate and MAP. This is likely a result confounded with the high degree of missingness in these two vitals.

	hemoglobin	heart rate	MAP	lactate
state 1 ($\mathbf{A}_1$)	0.142[0.048, 0.265]	0.969[0.966, 0.972]	0.880[0.873, 0.887]	0.863[0.838, 0.884]
state 2 ($\mathbf{A}_2$)	0.197[0.062, 0.385]	0.004[0.002, 0.008]	0.002[0.001, 0.003]	0.775[0.688, 0.848]
state 3 ($\mathbf{A}_3$)	0.494[0.130, 0.824]	0.273[0.217, 0.299]	0.009[0.004, 0.018]	0.935[0.863, 0.978]
state 4 ($\mathbf{A}_4$)	0.274[0.085, 0.633]	0.001[0.001, 0.003]	0.001[0.001, 0.002]	0.772[0.642, 0.868]
state 5 ($\mathbf{A}_5$)	0.101[0.036, 0.198]	0.364[0.339, 0.387]	0.143[0.098, 0.171]	0.260[0.122, 0.383]

Table 5: The posterior median estimates and 95% credible sets for the vector AR coefficients.

5.2 Case Study

Using the parameter estimates from Section 5.1, we implement our state-sampling routine on a test set comprised of five real patient encounters (Subjects A, B, C, D, and E). The MCMC routine is run for 2,000,000 iterations where the only parameters being sampled are the random effect coefficients $\boldsymbol{\alpha}_*^{(i)}$ and $\boldsymbol{\gamma}^{(i)}$, and the state sequences $\boldsymbol{b}^{(i)}$. All remaining parameters are fixed at their posterior medians. Using the information from the last 500,000 steps, Figure 5 presents the results for one out of the five test subjects (Subject C), and the remaining four test subject case studies are in Supplementary Materials Section 13.

5.2.1 Subject C

Figure 5 follows a 72-year-old woman who underwent resection and replacement of a thoracoabdominal aortic aneurysm. Her operative course was notable for significant blood loss and coagulopathy requiring multiple transfusions. During the first several hours of admission to the ICU post-operatively (shown in Figure 5), she had 2 episodes of increased chest tube output in the setting of coagulopathy (INR 2.4, fibrinogen 84, platelets 18) for which she received plasma, cryoprecipitate, platelets, and RBCs (only RBCs indicated in Figure 5). Over the next 12-16 hours, she developed refractory acidosis and progressive hemodynamic deterioration consistent with septic shock. A CT angiogram revealed a superior mesenteric artery (SMA) dissection, and she was rapidly taken back to the operating room. Unfortunately, half of her small bowel was found to be ischemic, and her colon was

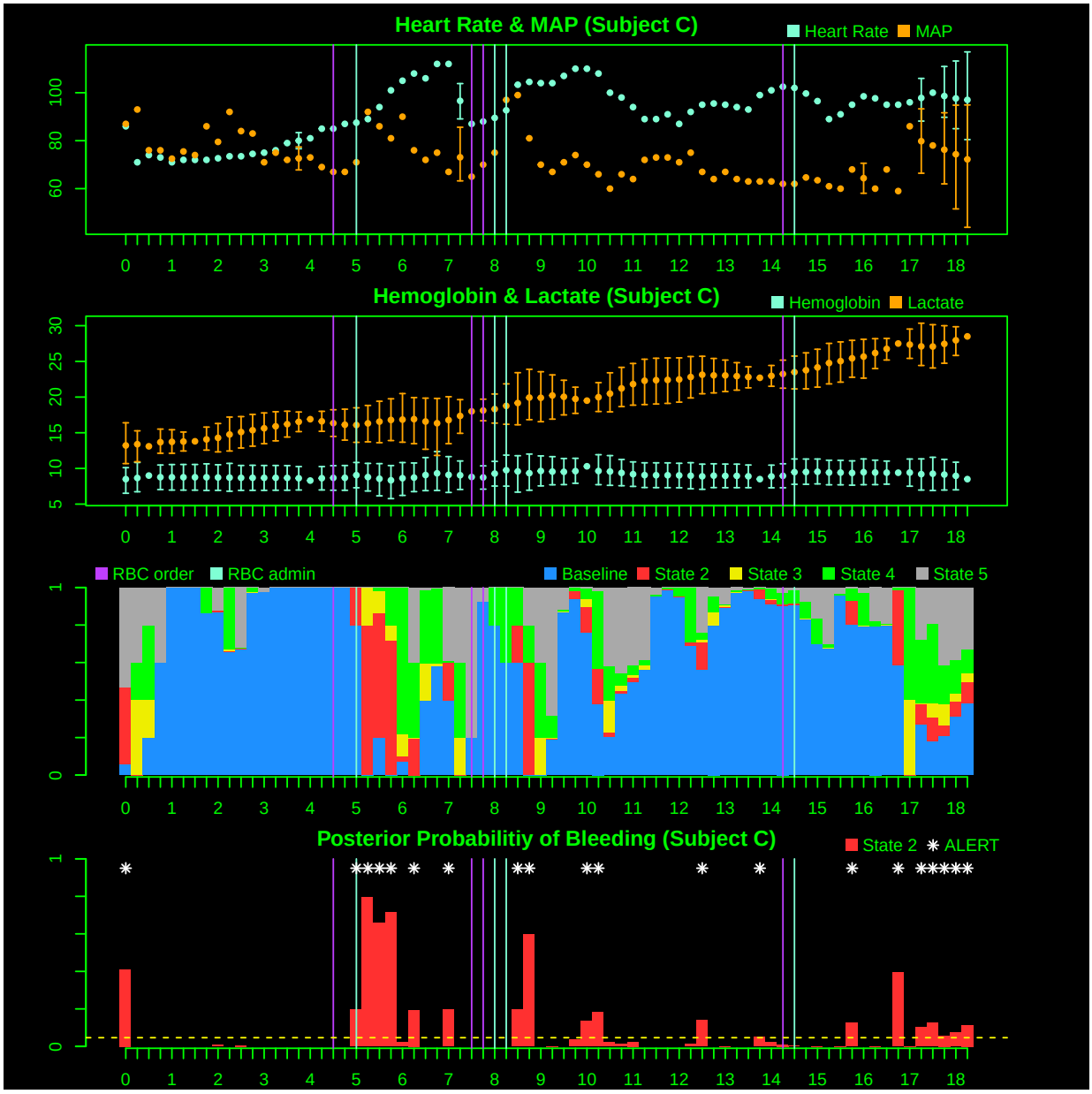

Figure 5: The top two panels correspond to the longitudinal vital measurements. The points with the error bars correspond to missing values; the error bars are empirical 95% credible intervals for the imputed response values. The third panel depicts the discrete posterior probability distributions of the latent states, at each time point. The bottom panel is the posterior probability of state 2 at each time point, and the yellow dashed line represents the threshold $\hat{c} = 0.0465$ determined from Section 4.2. The white stars indicate that the posterior probability of state 2 exceeds the threshold. The purple and turquoise vertical lines represent RBC transfusion order and administration times, respectively.

necrotic requiring a total colectomy. The SMA bypass graft was revised, and her abdomen was packed open. Blood loss was estimated at 350 mL. To optimize bowel perfusion, she was transfused with 2 additional units of RBCs. Despite aggressive care after returning to

the ICU, her shock state progressed with severe metabolic acidosis, escalating vasopressor and inotropic support and eventual myocardial dysfunction (an overall state of mixed shock including septic, hypovolemic due to capillary leak, and cardiogenic). Care was withdrawn.

Figure 5 appears to reflect the patient's dynamic course with temporal accuracy as reflected by the transition in states. Although the figure indicates state 2 suggesting bleeding on several occasions (as seen in red), these events were not always clinically correlated with hemorrhage. Instead, state 2 was suggested when septic and cardiogenic shock were present and progressing rapidly. This is likely due to the overlap in the presentation of different shock types in terms of the vital signs heart rate, MAP, and lactate.

6 Conclusion

The successful development of the RSM described in this manuscript will advance patient monitoring and medical interventions, particularly for transfusion medicine. Early detection of occult bleeding has been a persistent and difficult challenge for clinicians. The expected indicators of blood loss, including deterioration in vital signs and reductions in hemoglobin are typically delayed in their presentation. Much of this is due to the remarkable compensation of human physiology. It is only when compensation begins to fail that we see the typical changes in vital signs we associate with hemorrhage. These changes are well described by the American College of Surgeons Advanced Trauma Life Support (ATLS) classification of blood loss. Using this system, heart rate only starts to increase to >100 beats per minute when blood loss reaches 15-30% of total blood volume or 750-1500 mL, and MAP only begins to decline when blood loss reaches 30-40% of total blood volume, or 1500-2000mL (American College of Surgeons Committee on Trauma 2018). With the failure of physiological compensation, hypoperfusion, end-organ injury and death rapidly

ensue. Early detection and treatment of bleeding events has the potential to interrupt the downward spiral of hemorrhagic shock, prevent organ failure, reduce death, and limit total transfusion requirements. These implications are meaningful to patients and important from the standpoint of resource utilization. To make strides in this area, it is clear that we need tools and early warning systems that outperform clinicians monitoring basic vital signs and changes in hemoglobin. We hope that leveraging novel statistical frameworks will facilitate better identification and earlier intervention for this challenging problem.

The current manuscript presents a retrospective study on existing patient encounters; the immediate next step for future work is to produce an interpretable R Shiny Application with a user-interface tailored to clinical use in a real-time ICU setting. The goal is to develop an RSM that will quantify uncertainty about various medical conditions that are inherently difficult to identify in real-time, more than just internal bleeding. Additionally, the mean structure of our conditional response is clinically advisable, however, our approach requires assuming that the shock phenotype for an individual is the same over time. In other words, if a patient suffers from two distinct hemorrhage events during their ICU encounter, the current model assumes that the expected change in vitals per time instance will be the same for both bleeding events; physiologically, this need not be the case. A more flexible characterization of the latent states is worth exploring next.

7 Data Availability Statement

The data that support the findings of this study are available from the corresponding author [EBK]upon reasonable request.

8 Funding

Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R56HL155373.

References

Ailliot, P. & Monbet, V. (2012), 'Markov-switching autoregressive models for wind time series', *Environmental Modelling & Software* **30**, 92–101.

Albert, J. H. & Chib, S. (1993), 'Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts', *Journal of Business & Economic Statistics* **11**(1), 1–15.

Altman, R. M. (2007), 'Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting', *Journal of the American Statistical Association* **102**(477), 201–210.

American College of Surgeons Committee on Trauma (2018), *Advanced Trauma Life Support ®Student Manual*, 10th edn, American College of Surgeons, Chicago, IL.

Bedoya, A. D., Futoma, J., Clement, M. E., Corey, K., Brajer, N., Lin, A., Simons, M. G., Gao, M., Nichols, M., Balu, S. et al. (2020), 'Machine learning for early detection of sepsis: an internal and temporal validation study', *JAMIA open* **3**(2), 252–260.

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. PMID: 12529874.

Carter, C. K. & Kohn, R. (1994), 'On Gibbs sampling for state space models', *Biometrika* **81**(3), 541–553.

Clarke, J. R., Trooskin, S. Z., Doshi, P. J., Greenwald, L. & Mode, C. J. (2002), 'Time to laparotomy for intra-abdominal bleeding from trauma does affect survival for delays up to 90 minutes', *Journal of Trauma and Acute Care Surgery* **52**(3), 420–425. PMID: 11901314.

Convertino, V. A. & Cardin, S. (2022), 'Advanced medical monitoring for the battlefield: a review on clinical applicability of compensatory reserve measurements for early and accurate hemorrhage detection', *Journal of Trauma and Acute Care Surgery* **93**(2S), S147–S154.

Djuric, P. M. & Chun, J.-H. (2002), 'An MCMC sampling approach to estimation of nonstationary hidden Markov models', *IEEE Transactions on Signal Processing* **50**(5), 1113–1123.

Evans, L., Rhodes, A., Alhazzani, W., Antonelli, M., Coopersmith, C. M., French, C., Machado, F. R., Mcintyre, L., Ostermann, M., Prescott, H. C. et al. (2021), 'Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021', *Critical care medicine* **49**(11), e1063–e1143.

Fearnhead, P. & Sherlock, C. (2025), 'MCMC for state space models'.
URL: *https://arxiv.org/abs/2510.04932*

Holcomb, J. B., Tilley, B. C., Baraniuk, S., Fox, E. E., Wade, C. E., Podbielski, J. M., Del Junco, D. J., Brasel, K. J., Bulger, E. M., Callcut, R. A. et al. (2015), 'Transfusion of plasma, platelets, and red blood cells in a 1:1:1 vs a 1:1:2 ratio and mortality in patients with severe trauma: the PROPPR randomized clinical trial', *Journal of the American Medical Association* **313**(5), 471–482. PMID: 25647203.

Hooper, N. & Armstrong, T. J. (2022), Hemorrhagic shock, *in* 'StatPearls [internet]', StatPearls Publishing.

Hornbrook, M. C., Goshen, R., Choman, E., O'Keeffe-Rosetti, M., Kinar, Y., Liles, E. G. & Rust, K. C. (2017), 'Early colorectal cancer detected by machine learning model using gender, age, and complete blood count data', *Digestive Diseases and Sciences* **62**(10), 2719–2727.

Itzhak, N., Pessach, I. M. & Moskovitch, R. (2023), 'Prediction of acute hypertensive episodes in critically ill patients', *Artificial Intelligence in Medicine* **139**, 102525.

Jackson, C. H., Sharples, L. D., Thompson, S. G., Duffy, S. W. & Couto, E. (2003), 'Multistate Markov models for disease progression with classification error', *The Statistician* **52**(2), 193–209.

Jha, S. & Feng, W.-c. (2023), GRAPPEL: A graph-based approach for early risk assessment of acute hypertension in critical care, *in* 'Proceedings of the 14th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics', pp. 1–6.

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.

Kendall, E. B., Williams, J. P., Hermansen, G. H., Bois, F. & Thanh, V. H. (2025), 'Beyond time-homogeneity for continuous-time multistate Markov models', *Journal of Computational and Graphical Statistics* **34**(2), 668–682.

Kotwal, R. S., Scott, L. L., Janak, J. C., Tarpey, B. W., Howard, J. T., Mazuchowski, E. L., Butler, F. K., Shackelford, S. A., Gurney, J. M. & Stockinger, Z. T. (2018), 'The effect of prehospital transport time, injury severity, and blood transfusion on survival of US military casualties in Iraq', *Journal of Trauma and Acute Care Surgery* **85**(1S), S112–S121. PMID: 29334570.

Kwon, J.-M., Cho, Y., Jeon, K.-H., Cho, S., Kim, K.-H., Baek, S. D., Jeung, S., Park, J. & Oh, B.-H. (2020), 'A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study', *The Lancet Digital Health* **2**(7), e358–e367.

Langrock, R., Adam, T., Leos-Barajas, V., Mews, S., Miller, D. L. & Papastamatiou, Y. P. (2018), 'Spline-based nonparametric inference in general state-switching models', *Statistica Neerlandica* **72**(3), 179–200.

Latif, R. K., Clifford, S. P., Baker, J. A., Lenhardt, R., Haq, M. Z., Huang, J., Farah, I. & Businger, J. R. (2023), 'Traumatic hemorrhage and chain of survival', *Scandinavian journal of trauma, resuscitation and emergency medicine* **31**(1), 25.

Li-wei, H. L., Nemati, S. & Mark, R. G. (2015), Hemodynamic monitoring using switching autoregressive dynamics of multivariate vital sign time series, *in* '2015 Computing in Cardiology Conference (CinC)', IEEE, pp. 1065–1068.

Li, Y., Ning, S., Calvo, S. E., Mootha, V. K., Liu, J. S. et al. (2019), 'Bayesian hidden Markov tree models for clustering genes with shared evolutionary history', *Annals of Applied Statistics* **13**(1), 606–637.

Lu, J., Xiao, Q. & Bauer, C. (2023), 'A Bayesian circadian hidden Markov model to infer rest-activity rhythms using 24-hour actigraphy data', *arXiv preprint arXiv:2307.03832* .

Maier, C. L., Brohi, K., Curry, N., Juffermans, N. P., Mora Miquel, L., Neal, M. D., Shaz, B. H., Vlaar, A. P. & Helms, J. (2024), 'Contemporary management of major haemorrhage in critical care', *Intensive Care Medicine* **50**(3), 319–331.

McLouth, J., Elstrott, S., Chaibi, Y., Quenet, S., Chang, P. D., Chow, D. S. & Soun, J. E. (2021), 'Validation of a deep learning tool in the detection of intracranial hemorrhage and large vessel occlusion', *Frontiers in Neurology* **12**, 656112.

Moore, E. E., Moore, H. B., Kornblith, L. Z., Neal, M. D., Hoffman, M., Mutch, N. J., Schöchl, H., Hunt, B. J. & Sauaia, A. (2021), 'Trauma-induced coagulopathy', *Nature Reviews Disease Primers* **7**(1), 30.

Pannu, H. S., Ahuja, S., Dang, N., Soni, S. & Malhi, A. K. (2020), 'Deep learning based image classification for intestinal hemorrhage', *Multimedia Tools and Applications* **79**(29), 21941–21966.

Peltan, I. D., Brown, S. M., Bledsoe, J. R., Sorensen, J., Samore, M. H., Allen, T. L. & Hough, C. L. (2019), 'ED door-to-antibiotic time and long-term mortality in sepsis', *Chest* **155**(5), 938–946. PMID: 30779916.

Puerto-Santana, C., Larrañaga, P. & Bielza, C. (2021), 'Autoregressive asymmetric linear Gaussian hidden Markov models', *IEEE transactions on pattern analysis and machine intelligence* **44**(9), 4642–4658.

Rabiner, L. R. (1989), 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceedings of the IEEE* **77**(2), 257–286.

Robert, C. P., Celeux, G. & Diebolt, J. (1993), 'Bayesian estimation of hidden Markov chains: a stochastic implementation', *Statistics & Probability Letters* **16**(1), 77–83.

Robert, C. P. & Titterington, D. (1998), 'Reparameterization strategies for hidden Markov models and Bayesian approaches to maximum likelihood estimation', *Statistics and Computing* **8**, 145–158.

Sampalis, J. S., Denis, R., Lavoie, A., Frechette, P., Boukas, S., Nikolis, A., Benoit, D., Fleiszer, D., Brown, R., Churchill-Smith, M. et al. (1999), 'Trauma care regionalization: a process-outcome evaluation', *Journal of Trauma and Acute Care Surgery* **46**(4), 565–581. PMID: 10217218.

Sampalis, J. S., Lavoie, A., Williams, J., Mulder, D. S. & Kalina, M. (1993), 'Impact of on-site care, prehospital time, and level of in-hospital care on survival in severely injured patients', *Journal of Trauma* **34**(2), 252–261. PMID: 8459466.

Satten, G. A. & Longini, I. M. (1996), 'Markov chains with measurement error: Estimating the 'true' course of a marker of the progression of human immunodeficiency virus disease', *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **45**(3), 275–309.

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.

Scott, S. L., James, G. M. & Sugar, C. A. (2005), 'Hidden Markov models for longitudinal comparisons', *Journal of the American Statistical Association* **100**(470), 359–369.

Shirley, K. E., Small, D. S., Lynch, K. G., Maisto, S. A. & Oslin, D. W. (2010), 'Hidden Markov models for alcoholism treatment trial data', *Annals of Applied Statistics* **4**(1), 366–395.

Sidrow, E., Heckman, N., Fortune, S. M., Trites, A. W., Murphy, I. & Auger-Méthé, M. (2022), 'Modelling multi-scale, state-switching functional data with hidden Markov models', *Canadian Journal of Statistics* **50**(1), 327–356.

Stanculescu, I., Williams, C. K. & Freer, Y. (2013), 'Autoregressive hidden Markov models for the early detection of neonatal sepsis', *IEEE Journal of Biomedical and Health Informatics* **18**(5), 1560–1570. PMID: 25192568.

Stensballe, J., Henriksen, H. H. & Johansson, P. I. (2017), 'Early haemorrhage control and management of trauma-induced coagulopathy: the importance of goal-directed therapy', *Current Opinion in Critical Care* **23**(6), 503–510. PMID: 29059118.

Storlie, C., Sexton, J., Pakin, S., Lang, M., Reich, B. & Rust, W. (2014), 'Modeling and predicting power consumption of high performance computing jobs', *arXiv preprint arXiv:1412.5247* .

Strehlow, M. C. (2010), 'Early identification of shock in critically ill patients', *Emergency Medicine Clinics* **28**(1), 57–66.

Trabelsi, D., Mohammed, S., Chamroukhi, F., Oukhellou, L. & Amirat, Y. (2013), 'An unsupervised approach for automatic activity recognition based on hidden Markov model regression', *IEEE Transactions on Automation Science and Engineering* **10**(3), 829–835.

Tran, A., Taljaard, M., Abdulaziz, K. E., Matar, M., Lampron, J., Steyerberg, E. W. & Vaillancourt, C. (2020), 'Early identification of the need for major intervention in patients with traumatic hemorrhage: development and internal validation of a simple bleeding score', *Canadian Journal of Surgery* **63**(5), E422.

Triantafyllopoulos, K. et al. (2021), *Bayesian inference of state space models*, Springer.

Turek, D., de Valpine, P. & Paciorek, C. J. (2016), 'Efficient Markov chain Monte Carlo sampling for hierarchical hidden Markov models', *Environmental and ecological statistics* **23**, 549–564.

Volpe, V. V., Kendall, E. B., Collins, A. N., Graham, M. G., Williams, J. P. & Holochwost, S. J. (2025), 'Prior exposure to racial discrimination and patterns of acute parasympathetic nervous system responses to a race-related stress task among Black adults', *Psychophysiology* **62**(1), e14713.

Wang, E. T., Chiang, S., Haneef, Z., Rao, V. R., Moss, R. & Vannucci, M. (2023), 'Bayesian nonhomogeneous hidden Markov model with variable selection for investigating drivers of seizure risk cycling', *Annals of Applied Statistics* **17**(1), 333–356.

Williams, J. P., Hermansen, G. H., Strand, H., Clayton, G. & Nygård, H. M. (2024), 'Bayesian hidden Markov models for latent variable labeling assignments in conflict research: Application to the role ceasefires play in conflict dynamics', *Annals of Applied Statistics* **18**(3), 2034–2061.

Williams, J. P., Storlie, C. B., Therneau, T. M., Jack Jr, C. R. & Hannig, J. (2020), 'A Bayesian approach to multistate hidden Markov models: Application to dementia progression', *Journal of the American Statistical Association* **115**(529), 16–31.

Zhao, T., Wang, Z., Cumberworth, A., Gsponer, J., de Freitas, N. & Bouchard-Côté, A. (2016), 'Bayesian analysis of continuous time Markov chains with application to phylogenetic modeling', *Bayesian Analysis* **11**(4), 1203–1237.