### Title

A Bayesian multilevel time-varying framework for joint modeling of hospitalization and survival in patients on dialysis.




### Journal

Statistics in Medicine, 41(29)



### Authors

Kürüm, Esra
Nguyen, Danh
Banerjee, Sudipto
et al.



### Publication Date

2022-12-20

### DOI

10.1002/sim.9582

Peer reviewed

Published in final edited form as:

*Stat Med*. 2022 December 20; 41(29): 5597–5611. doi:10.1002/sim.9582.

# A Bayesian multilevel time-varying framework for joint modeling of hospitalization and survival in patients on dialysis

**Esra Kürüm**[*,1], **Danh V. Nguyen**[2], **Sudipto Banerjee**[3], **Yihao Li**[3], **Connie M. Rhee**[2,4], **Damla Şentürk**[3]

[1]Department of Statistics, University of California, Riverside, CA, USA

[2]Department of Medicine, University of California Irvine, CA, USA

[3]Department of Biostatistics, University of California, Los Angeles, CA, USA

[4]Harold Simmons Center for Chronic Disease Research and Epidemiology, University of California Irvine School of Medicine, CA, USA

## Abstract

Over 782,000 individuals in the U.S. have end-stage kidney disease with about 72% of patients on dialysis, a life-sustaining treatment. Dialysis patients experience high mortality and frequent hospitalizations, at about twice per year. These poor outcomes are exacerbated at key time periods, such as the fragile period after transition to dialysis. In order to study the time-varying effects of modifiable patient and dialysis facility risk factors on hospitalization and mortality, we propose a novel Bayesian multilevel time-varying joint model. Efficient estimation and inference is achieved within the Bayesian framework using Markov Chain Monte Carlo, where multilevel (patient- and dialysis facility-level) varying coefficient functions are targeted via Bayesian P-splines. Applications to the United States Renal Data System, a national database which contains data on nearly all patients on dialysis in the U.S., highlight significant time-varying effects of patient- and facility-level risk factors on hospitalization risk and mortality. Finite sample performance of the proposed methodology is studied through simulations.

### Keywords

End-stage kidney disease; Markov Chain Monte Carlo; Mixed-effects models; Varying-coefficient models; United States Renal Data System

## 1 | INTRODUCTION

End-stage kidney disease (ESKD) affects more than 782,000 individuals in the United States, with about 72% of patients on dialysis, a life-sustaining treatment[1]. Patients on dialysis experience a high burden of complex comorbid conditions and hence are frequently hospitalized, about twice a year, where these poor outcomes are exacerbated at key time periods after transitioning to dialysis. Frequent hospitalizations are a major contributor

[*]**Correspondence:** Esra Kürüm, Department of Statistics, University of California, Riverside, CA, USA., esra.kurum@ucr.edu.

to mortality in the dialysis population, in whom the mortality risk is significantly higher compared to other Medicare populations, including those with cancer and cardiovascular disease. Therefore, an important goal is to identify modifiable patient and dialysis facility risk factors and their time-dynamic effects on the correlated outcomes of hospitalization risk and mortality to guide patient care strategies.

To understand the time-varying effects of multilevel risk factors, we analyze data from the United States Renal Data System (USRDS), a large national database, where the data structure is hierarchical with longitudinal hospitalizations over time, nested within subjects, and subjects further nested within dialysis facilities where they receive regular care. The data also contains potential risk factors at both the patient (e.g., baseline demographics and comorbidities) and dialysis facility levels (e.g., dialysis staffing level, which can contribute to hospital readmission[2]). Thus, our modeling approach needs to take this multilevel structure into account while studying the impact of risk factors from both levels of the hierarchy. In addition, ESKD patients typically stay on dialysis for the duration of their lives or until successful kidney transplantation. As the patients' clinical conditions may change after transitioning to dialysis, their hospitalization and mortality risks also vary post-dialysis transition. Previous works documented these temporal changes in hospitalization risk and studied the time-varying effects of risk factors on hospitalizations in patients on dialysis[3,4,5,6]. However, time-varying effects of multilevel risk factors have not been studied for the correlated outcomes of hospitalizations and mortality, jointly.

There is extensive literature on joint modeling of longitudinal outcomes and survival as these models have become a valuable tool in analyzing mixed outcomes. See[7] and[8] for a detailed review. A popular approach to joint modeling is the shared-parameter modeling framework, where dependence between the outcomes is modeled via a set of shared underlying random effects[9,10,11,12,13,14,15]. While shared random effects models have been popular for a two-level hierarchy, that is, repeated measurements nested within subjects, there is a paucity of works which consider three-level hierarchical data, with longitudinal data nested in subjects and subjects nested in a higher clustering units such as dialysis facilities, understandably due to significant computational challenges[16,17,18]. Liu et al. (2008) considered joint modeling of a continuous longitudinal outcome with survival in a frequentist framework utilizing an Expectation-Maximization (EM) algorithm with Gauss quadrature for integration of the random effects (REs), and Brilleman et al. (2019) considered joint modeling in a three-level hierarchy using a Bayesian framework. However, both works considered only subject-level risk factors. Kurum et al. (2021) extended the frequentist approach to model effects of multilevel risk factors (at both the subject and facility levels) on the longitudinal outcome and survival by utilizing an EM algorithm incorporating fully exponential Laplace approximations for integration of the potentially high-dimensional multilevel REs in the three-level hierarchical setting. However, under the EM algorithm, the likelihood-based standard errors (SEs) lead to bias in estimation of the true SEs, necessitating bootstrap inference in the frequentist setting[12,13,14,18] and increasing the already high computational burden.

Furthermore, none of the aforementioned joint modeling approaches for three-level hierarchical data consider modeling time-varying effects. More specifically, only few works

in the joint modeling literature allow for time-varying effects, either time-varying response-predictor relationships or time-varying relationships between the outcomes modeled[19,20,21]. However, these methods only allow for a two-level hierarchy, that is, repeated measurements nested within subjects. Our proposed Bayesian multilevel time-varying joint model (BMT-JM) accounts for time-varying effects of multilevel risk factors on both hospitalization and mortality as well as the time-varying relationship between the two outcomes, while also accommodating the three-level hierarchical data structure of the USRDS data through multilevel REs. To study the time-varying effects of multilevel covariates and the time-varying relationships between the outcomes, we employ varying-coefficient models[22,23,24,3], which are widely used to study time-varying effects in longitudinal studies. In our proposed approach, the time-varying coefficients are estimated via Bayesian P-splines[25], leading to a computationally efficient algorithm. In addition, valid inference is achieved within the Bayesian framework based on Markov Chain Monte Carlo (MCMC), avoiding reliance on computationally intensive bootstrap procedures.

Thus, our paper makes a novel contribution to the joint modeling literature by proposing an approach that can handle complex hierarchical data structures and feasibly estimate all manners of time-varying relationships (response-predictor and response-response). Furthermore, due to the Bayesian estimation framework, our method provides a computationally efficient algorithm for estimation and inference, overcoming the drawbacks of the standard estimation methods based on the EM algorithm. The remainder of the paper is organized as follows. The proposed BMT-JM and the Bayesian estimation, along with the inference procedure, are described in Section 2. Simulation studies to examine the efficacy of estimation are given in Section 3. In Section 4, we illustrate the proposed BMT-JM to jointly model longitudinal hospitalization risk and survival using the USRDS data. We conclude with a brief discussion in Section 5.

## 2 | BAYESIAN MULTILEVEL TIME-VARYING JOINT MODELS

### 2.1 | Model specification

In our joint modeling framework, we start with defining the submodels for each outcome. For the longitudinal submodel, we let $Y_{ij}(t)$ represent the longitudinal outcome for the $j$th subject (patient), $j = 1, \ldots, n_i$, at the $i$th cluster (facility), $i = 1, \ldots, n$, at time $t$. In our data application, the time index $t$ is taken to be time (days) starting from when a patient transition to dialysis. Also, in our data application $Y_{ij}(t)$ is a binary longitudinal outcome defined as the indicator of at least one hospitalization in a 3-month follow-up window with midpoint $t$ for the subject $j$ at facility $i$. Let $\mathbf{X}_{ij} = \left(X_{ij1}, \ldots, X_{ijp}\right)^{\mathrm{T}}$ and $\mathbf{Z}_{i(j)} = \left\{Z_{i(j)1}, \ldots, Z_{i(j)q}\right\}^{\mathrm{T}}$ be the subject- and facility-level predictors with the corresponding time-varying coefficients $\boldsymbol{\beta}_X(t) = \left\{\beta_{X_1}(t), \ldots, \beta_{X_p}(t)\right\}^{\mathrm{T}}$ and $\boldsymbol{\beta}_Z(t) = \left\{\beta_{Z_1}(t), \ldots, \beta_{Z_q}(t)\right\}^{\mathrm{T}}$, respectively. For the USRDS data, the facility-level characteristics, such as nurse-to-patient ratio, are reported annually, and $\mathbf{Z}_{i(j)}$ denote those characteristics recorded in the calendar year prior to the $j$th patient initiating dialysis. Hence, these predictors are indexed not only by the dialysis facility index $i$ but also by the subject index $j$. In order to capture the time-dynamic trends in the data, we propose a generalized multilevel varying-coefficient model for the longitudinal submodel

$$m_{ij}(t) = E\left\{Y_{ij}(t) \mid \mathbf{X}_{ij}, \mathbf{Z}_{i(j)}, b_{ij}, b_i\right\} = g^{-1}\left\{\mathbf{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_X(t) + \mathbf{Z}_{i(j)}^{\mathrm{T}}\boldsymbol{\beta}_Z(t) + b_{ij} + b_i\right\},$$

where $g(\cdot)$ denotes the canonical link, and $b_{ij}$ and $b_i$ are the subject- and facility-level REs, respectively. For our motivating data application, the link function takes the form of the logit link $g(p) = log\{p/(1-p)\}$. The subject- and facility-level REs are assumed to be independent, and each follows a normal distribution such that $b_{ij} \sim N\left(0, \sigma_S^2\right)$ and $b_i \sim N\left(0, \sigma_F^2\right)$. Although we make distributional assumptions about the REs, empirical results show that the parameter estimation and inference in joint modeling are robust to these specifications, similar to standard joint models[12,13,14].

In the survival submodel, the true and observed event (death) times are denoted by $T_{ij}^*$ and $T_{ij}$, respectively, for subject $j$ at cluster $i$. The observed event time is defined as the minimum of the potential censoring time $C_{ij}$ and $T_{ij}^*$. The event indicator is defined as $\delta_{ij} = I\left(T_{ij} \le C_{ij}\right)$, where $I(\cdot)$ is the indicator function. We assume that the censoring mechanism is noninformative; that is, it is independent of the longitudinal process and the covariates. The proposed submodel for the survival outcome is the Cox model with time-varying coefficients, where the hazard of death at time $t$, accounting for the hospitalization history up to time $t$ is defined as

$$\begin{aligned} h_{ij}\left\{t \mid \mathcal{M}_{ij}(t), \mathbf{X}_{ij}, \mathbf{Z}_{i(j)}\right\} &= \lim_{\Delta_t \to 0} \Pr\left\{t \le T_{ij}^* < t + \Delta_t \mid T_{ij}^* \ge t, \mathcal{M}_{ij}(t), \mathbf{X}_{ij}, \mathbf{Z}_{i(j)}\right\} \\ &= h_0(t)\exp\left\{\mathbf{X}_{ij}^{\mathrm{T}}\boldsymbol{\gamma}_X(t) + \mathbf{Z}_{i(j)}^{\mathrm{T}}\boldsymbol{\gamma}_Z(t) + \alpha(t)m_{ij}(t)\right\}, \end{aligned}$$

with $\mathcal{M}_{ij}(t) = \left\{m_{ij}(u), 0 \le u < t\right\}$ denoting the history of the true unobserved longitudinal process up to time point $t$, $\boldsymbol{\gamma}_X(t) = \left\{\gamma_{X_1}(t), \ldots, \gamma_{X_p}(t)\right\}^{\mathrm{T}}$ and $\boldsymbol{\gamma}_Z(t) = \left\{\gamma_{Z_1}(t), \ldots, \gamma_{Z_q}(t)\right\}^{\mathrm{T}}$ denoting the multilevel covariate effects on survival, $h_0(t)$ denoting the baseline hazard and $\alpha(t)$ denoting the regression coefficient function that quantifies the time-varying effect of the longitudinal outcome on the risk of death. The corresponding survival function is

$$\begin{aligned} \mathcal{S}_{ij}\left\{t \mid \mathcal{M}_{ij}(t), \mathbf{X}_{ij}, \mathbf{Z}_{i(j)}\right\} &= Pr\left\{T_{ij}^* > t \mid \mathcal{M}_{ij}(t), \mathbf{X}_{ij}, \mathbf{Z}_{i(j)}\right\} \\ &= \exp\left[-\int_0^t h_0(u)\exp\left\{\mathbf{X}_{ij}^{\mathrm{T}}\gamma_X(u) + \mathbf{Z}_{i(j)}^{\mathrm{T}}\gamma_Z(u) + \alpha(u)m_{ij}(u)\right\}du\right]. \end{aligned}$$

In our data application, the definitions of the hazard and survival functions imply that the survival function depends on the entire history of the longitudinal outcome $\left(\mathcal{M}_{ij}(t)\right)$, that is, hospitalization risk up to time $t$, whereas the hazard of death at time $t$ depends only on the current value of the hospitalization risk score, $m_{ij}(t)$.

## 2.2 | Estimation

To estimate the parameters in our joint modeling framework, we propose a Bayesian estimation procedure and derive posterior inferences using an MCMC algorithm. Let $\theta(t) = \left\{\theta_y^{\mathrm{T}}(t), \theta_s^{\mathrm{T}}(t)\right\}^{\mathrm{T}}$, where $\theta_y(t) = \left\{\beta_X^{\mathrm{T}}(t), \beta_Z^{\mathrm{T}}(t), \sigma_S^2, \sigma_F^2\right\}^{\mathrm{T}}$ and $\theta_s(t) = \left\{\gamma_X^{\mathrm{T}}(t), \gamma_Z^{\mathrm{T}}(t), \alpha(t), \theta_{h_0}\right\}^{\mathrm{T}}$ are the parameters in the longitudinal and survival submodels, respectively, with $\theta_{h_0}$ denoting the coefficients used to model the baseline hazard. We utilize Bayesian P-splines in estimation of the baseline hazard as well as the multilevel varying coefficient functions as outlined below. The joint likelihood is obtained under the conditional independence assumption, that is, the REs account for the association between the two outcomes, and given the REs, the outcomes are independent. Furthermore, we assume that, in addition to the time-varying effects, the subject-level REs also contribute to modeling the correlation between the longitudinal measurements within a subject, leading to the conditional likelihood

$$p\left\{T_{ij}, \delta_{ij}, \mathbf{Y}_{ij} \mid b_{ij}, b_i; \theta(t)\right\} = p\left\{T_{ij}, \delta_{ij} \mid b_{ij}, b_i; \theta(t)\right\} p\left\{\mathbf{Y}_{ij} \mid b_{ij}, b_i; \theta(t)\right\},$$

$$p\left\{\mathbf{Y}_{ij} \mid b_{ij}, b_i; \theta(t)\right\} = \prod_{k=1}^{n_{ij}} p\left\{Y_{ijk} \mid b_{ij}, b_i; \theta_y(t)\right\},$$

where for the $j$th patient at the $i$th facility, $\mathbf{Y}_{ij} = \left\{Y_{ij1}, \ldots, Y_{ijn_{ij}}\right\}^{\mathrm{T}}$ denotes the $n_{ij} \times 1$ vector of longitudinal outcomes with $Y_{ijk} = Y_{ij}(t_{ijk})$, $k = 1, \ldots, n_{ij}$, and $t_{ijk}$ denoting the midpoint of the $k$th three-month interval in the follow-up period. Thus, the posterior distribution

$$\begin{aligned} p\left\{\theta(t), b_{ij}, b_i \mid \mathbf{Y}_{ij}, T_{ij}, \delta_{ij}\right\} &\propto p\left\{T_{ij}, \delta_{ij}, \mathbf{Y}_{ij} \mid b_{ij}, b_i, \theta(t)\right\} p\left\{b_{ij}, b_i, \theta(t)\right\} \\ &\propto \prod_{k=1}^{n_{ij}} p\left\{Y_{ijk} \mid b_{ij}, b_i, \theta_y(t)\right\} p\left\{T_{ij}, \delta_{ij} \mid b_{ij}, b_i, \theta(t)\right\} p \\ &\left\{b_{ij}, b_i \mid \theta(t)\right\} p\left\{\theta(t)\right\}. \end{aligned} \tag{1}$$

In our data application, where the longitudinal outcome is binary, the likelihood contribution from the longitudinal submodel is given by

$$p\left\{\mathbf{Y}_{ij} \mid b_{ij}, b_i, \theta_y(t)\right\} = \prod_{k=1}^{n_{ij}} \frac{\exp\left[\left\{\mathbf{X}_{ij}^{\mathrm{T}} \beta_X(t_{ijk}) + \mathbf{Z}_{i(j)}^{\mathrm{T}} \beta_Z(t_{ijk}) + b_{ij} + b_i\right\} Y_{ijk}\right]}{1 + \exp\left\{\mathbf{X}_{ij}^{\mathrm{T}} \beta_X(t_{ijk}) + \mathbf{Z}_{i(j)}^{\mathrm{T}} \beta_Z(t_{ijk}) + b_{ij} + b_i\right\}}. \tag{2}$$

In addition, the likelihood contribution of the survival submodel takes the form

$$
\begin{aligned}
p\{T_{ij}, \delta_{ij} \mid \mathcal{M}_{ij}(t), \theta(t)\} &= h_{ij}\{T_{ij} \mid \mathcal{M}_{ij}(T_{ij}), \theta(t)\}^{\delta_{ij}} \mathcal{S}_{ij}\{T_{ij} \mid \mathcal{M}_{ij}(T_{ij}), \theta(t)\} \\
&= \left[h_0(T_{ij}) \exp\left\{\mathbf{X}_{ij}^{\mathrm{T}} \boldsymbol{\gamma}_X(T_{ij}) + \mathbf{Z}_{i(j)}^{\mathrm{T}} \boldsymbol{\gamma}_Z(T_{ij}) + \alpha(T_{ij}) m_{ij}(T_{ij})\right\}\right]^{\delta_{ij}} \\
&\times \exp\left[-\int_0^{T_{ij}} h_0(u) \exp\left\{\mathbf{X}_{ij}^{\mathrm{T}} \boldsymbol{\gamma}_X(u) + \mathbf{Z}_{i(j)}^{\mathrm{T}} \boldsymbol{\gamma}_Z(u) + \alpha(u) m_{ij}(u)\right\} du\right]
\end{aligned}
\tag{3}
$$

.

The integral in the survival function does not have a closed-form solution; hence a numerical approximation is employed. We use the Gauss-Kronrod method.

We utilize Bayesian P-splines in estimation of the time-varying coefficients in our submodels[25]. In this approach, we use a relatively large number of equally spaced knots, and to avoid overfitting and obtain sufficiently smooth fitted curves, we apply a roughness penalty[26]. Specifically, the time-varying coefficient functions take the form

$$
\begin{aligned}
&\boldsymbol{\beta}_{X_\omega}(t) = \sum_{r=1}^{R} \phi_{X_\omega, r} B_r(t), \quad \boldsymbol{\beta}_{Z_v}(t) = \sum_{r=1}^{R} \phi_{Z_v, r} B_r(t), \\
&\boldsymbol{\gamma}_{X_\omega}(t) = \sum_{r=1}^{R} \psi_{X_\omega, r} B_r(t), \quad \boldsymbol{\gamma}_{Z_v}(t) = \sum_{r=1}^{R} \psi_{Z_v, r} B_r(t), \quad \alpha(t) \\
&= \sum_{r=1}^{R} \psi_{\alpha, r} B_r(t),
\end{aligned}
\tag{4}
$$

where $B_r(t)$ is the $r$th basis function of a B-spline, $\phi_{X_\omega}$, $\phi_{Z_v}$, $\psi_{X_\omega}$, $\psi_{Z_v}$, and $\boldsymbol{\psi}_\alpha$ are the R-dimensional vectors of spline coefficients, $\omega = 1, \ldots, p$, and $v = 1, \ldots q$. Moreover, we model the baseline hazard function $h_0(t)$ using the same flexible P-splines approach. More specifically, $\log\{h_0(t)\} = \sum_{r=1}^{R} \psi_{h_0, r} B_r(t)$ with $\psi_{h_0}$ denoting the corresponding vector of spline coefficients. Note that proposed estimation and inference procedures can also accommodate parametric and other nonparametric forms for the baseline hazard function.

Under the P-splines approach, the posterior density in (1) is rewritten as

$$
p\left(\theta_{\mathcal{P}}, b_{ij}, b_i \mid \mathbf{Y}_{ij}, T_{ij}, \delta_{ij}\right) \propto \prod_{k=1}^{n_{ij}} p\left(Y_{ijk} \mid b_{ij}, b_i, \theta_{y\mathcal{P}}\right) p\left(T_{ij}, \delta_{ij} \mid b_{ij}, b_i, \theta_{\mathcal{P}}\right) p\left(b_{ij}, b_i \mid \theta_{\mathcal{P}}\right) p(\theta_{\mathcal{P}}),
$$

where $\theta_{\mathcal{P}} = \left(\theta_{y\mathcal{P}}^{\mathrm{T}}, \theta_{s\mathcal{P}}^{\mathrm{T}}\right)^{\mathrm{T}}$ denotes the vector of parameters which includes the coefficients from the P-spline expansions with $\theta_{y\mathcal{P}} = \left(\phi_{X_w}^{\mathrm{T}}, \phi_{Z_v}^{\mathrm{T}}, \sigma_S^2, \sigma_F^2\right)^{\mathrm{T}}$ and $\theta_{s\mathcal{P}} = \left(\psi_{X_\omega}^{\mathrm{T}}, \psi_{Z_v}^{\mathrm{T}}, \psi_\alpha^{\mathrm{T}}, \psi_{h_0}^{\mathrm{T}}\right)^{\mathrm{T}}$. In terms of prior distributions, we use the normal distribution for the spline coefficients such that $\phi_{X_\omega} | \tau_{X_\omega} \sim N\left(0, \tau_{X_\omega} \mathcal{P}_{X_\omega}\right)$, $\phi_{Z_v} | \tau_{Z_v} \sim N\left(0, \tau_{Z_v} \mathcal{P}_{Z_v}\right)$,

$\psi_{X_\omega} \mid \kappa_{X_\omega} \sim N\left(0, \kappa_{X_\omega}\mathscr{P}_{X_\omega}\right)$, $\psi_{Z_\upsilon} \mid \kappa_{Z_\upsilon} \sim N\left(0, \kappa_{Z_\upsilon}\mathscr{P}_{Z_\upsilon}\right)$, $\psi_\alpha \mid \kappa_\alpha \sim N(0, \kappa_\alpha \mathscr{P}_\alpha)$, and $\psi_{h_0} \mid \kappa_{h_0} \sim N\left(0, \kappa_{h_0}\mathscr{P}_{h_0}\right)$ with $\omega = 1, \ldots p$ and $\upsilon = 1, \ldots, q$. The penalty matrix $\mathscr{P}_*$ ($*$denoting $X_\omega$, $Z_\upsilon$, $\alpha$, or $h_0$) in the above specified priors is calculated using the $\upsilon$th order difference matrix $\mathscr{D}_\upsilon$ of dimension $R \times R$, $\mathscr{P}_* = \mathscr{D}_\upsilon^{\mathrm{T}}\mathscr{D}_\upsilon + 10^{-6}\mathbf{I}$, where $\mathbf{I}$ denotes the $R \times R$ identity matrix. In applications, we utilize the commonly used second order difference matrix[26]. The variance parameters $\tau_*$ and $\kappa_*$, utilized in the priors of the longitudinal and survival submodel parameters, respectively, control the smoothness of the varying coefficient functions, where they are assigned Inverse Gamma (IG) priors, $\tau_* \sim IG\left(a_{1\tau_*}, a_{2\tau_*}\right)$ and $\kappa_* \sim IG\left(a_{1\kappa_*}, a_{2\kappa_*}\right)$. For the subject- and facility-level random effect variances, we also assume IG priors, $\sigma_S^2 \sim IG(a_{1S}, a_{2S})$ and $\sigma_F^2 \sim IG(a_{1F}, a_{2F})$. In the IG priors, the common practice is to set $a_1$'s equal to 1 and assign a small number to $a_2$'s. However, alternative specifications are also available, and especially for the hyperpriors $\tau_*$ and $\kappa_*$, where the selection depends on the smoothness of the function that is approximated via the P-splines approach[27]. As pointed out by[25] and[28], in some situations, the estimated functions may considerably depend on the particular choice of hyperparameters; in those cases, it is highly recommended to inspect the estimated results under a number of different choices for $a_1$'s and $a_2$'s.

For inference on the varying-coefficient functions, we use pointwise credible intervals and simultaneous credible bands[29]. Let $f(t)$ denote a single varying-coefficient function (taken to be $\alpha$(t), $h_0(t)$ or a varying coefficient function from the longitudinal or survival submodels, $\beta_*(t)$ or $\gamma_*(t)$, respectively) observed at time points $t_k$, for $k = 1, \ldots, K$. Let $\hat{f}(t)$ and SD$\{f(t)\}$ denote the mean and standard deviation of $f(t)$ obtained based on a total of $L$ MCMC samples, respectively. Then the $(1-\alpha)$ pointwise credible intervals are given by $\left[\hat{f}(t_k) \pm \Phi_{\alpha/2}\mathrm{SD}\{f(t_k)\}\right]$, where $\Phi_{\alpha/2}$ denotes the $100 \times (1 - \alpha/2)$-percentile of the standard normal distribution. For the simultaneous credible bands, let $c_\alpha$ be the $(1-\alpha)$ sample quantile of $\max_{k=1,\ldots,K} \mid f^{(\ell)}(t_k) - \hat{f}(t_k) \mid / SD\{f(t_k)\} \mid$ with $f^{(\ell)}(t)$, $\ell = 1\ldots, L$ denoting the $\ell$th MCMC sample. Then the $(1 - \alpha)$ simultaneous credible band for $f(t)$ is given as $\left[\hat{f}(t_k) \pm c_\alpha \mathrm{SD}\{f(t_k)\}\right]$.

We make three important remarks about our approach: (a) For simplicity of exposition, we describe the submodels using a common set of subject-and facility-level predictors; however, the estimation and inference procedures can accommodate design vectors with different dimensionality and composition. (b) Although the subject- and facility-level predictors are denoted as time-invariant covariates, they can include both baseline and time-varying covariates, under the condition that the time-varying covariates in the survival submodel are exogenous following the requirements of a proper Cox model. (c) In the estimation procedure, we use the same number of knots $R$ in equation (4) only for ease of notation, our estimation procedure can handle a different number of knots for each function.

All computations have been performed in R (version 4.0.2), and as there are no closed-form solutions for the posterior distributions, we fit our model using the Bayesian software JAGS

(version 4.3.0) via the rjags package[30]. All R codes and documentation for fitting the proposed BMT-MJM are made publicly available at https://github.com/esrakurum/BMT-JM.

## 3 | SIMULATION STUDIES

### 3.1 | Design

We examined the performance of the proposed estimation and inference procedures via simulation studies. We report the results of two simulation studies with $n = 200$ and $n = 500$ facilities. Motivated by the USRDS data application, the total number of patients within each facility was generated from a discrete uniform distribution on the interval [50, 162]. Similarly, mimicking the measurement times of the hospitalization outcome $Y_{ij}(\cdot)$ in the USRDS data, we assumed that the maximum number of repeated measurements per subject is 20, that is, every three months for a maximum of five years of follow-up. These repeated measurements were equally spaced on the interval [0, 1] before censoring by survival.

The subject-level covariates, $\mathbf{X}_{ij} = \left(X_{1ij}, X_{2ij}\right)^{\mathrm{T}}$, were simulated from normal distributions with means 0 and 1.5 and variances 1 and 0.5, respectively. The facility-level covariates, $\mathbf{Z}_{i(j)} = \left\{Z_{1i(j)}, Z_{2i(j)}\right\}^{\mathrm{T}}$, were simulated from normal distributions with means −0.3 and 1.5 and variances 1 and 0.5, respectively. The time-varying parameters in models (2) and (3) were generated as $\boldsymbol{\beta}_X(t) = \left\{\beta_{X_0}(t), \beta_{X_1}(t), \beta_{X_2}(t)\right\}^{\mathrm{T}} = \{cos(3/2\pi t) - 1, sin(2\pi t - 1/8), -sin(2\pi t - 1/8)\}^{\mathrm{T}}$, $\boldsymbol{\beta}_Z(t) = \left\{\beta_{Z_1}(t), \beta_{Z_2}(t)\right\}^{\mathrm{T}} = \{cos(\pi t - 0.5), -cos(\pi t - 0.5)\}^{\mathrm{T}}$, $\boldsymbol{\gamma}_X(t) = \left\{\gamma_{X_1}(t), \gamma_{X_2}(t)\right\}^{\mathrm{T}} = \{cos(2\pi t), -cos(2\pi t)\}^{\mathrm{T}}$, $\boldsymbol{\gamma}_Z = \left\{\gamma_{Z_1}(t), \gamma_{Z_2}(t)\right\}^{\mathrm{T}} = \{sin(3/4\pi t), -sin(3/4\pi t)\}^{\mathrm{T}}$, and $\alpha = sin(2\pi t)$. The Weibull function with $\lambda = 1.5$ was used to generate the baseline hazard $h_0(t)$. The subject- and facility-level REs were independently simulated from a normal distribution with mean zero and variances $\sigma_S^2 = 1.30$ and $\sigma_F^2 = 0.20$, respectively.

The longitudinal outcome $Y_{ij}(\cdot)$ was simulated using an underlying normal latent variable $Y_{ij}^*(\cdot)$ such that $Y_{ij}(\cdot) = I\left\{Y_{ij}^*(\cdot) > 0\right\}$, and the mean of $Y_{ij}^*(\cdot)$ was determined by the longitudinal submodel (2). For the survival submodel, the true event times, $T_{ij}^*$, for subjects were simulated using the inverse probability integral transformation with a Weibull baseline hazard function[31]. As described in Section 2.1, the observed time and the event indicator were calculated as $T_{ij} = \min\left(C_{ij}, T_{ij}^*\right)$ and $\delta_{ij} = I\left(T_{ij} \leq C_{ij}\right)$, respectively. Under this set-up, similar to the USRDS data, the overall hospitalization and censoring rates were approximately 29% and 62%, respectively.

### 3.2 | Simulation results

All time-varying coefficient functions and the baseline hazard function were estimated via the Bayesian P-splines with 20 equally spaced knots and a second-order penalty. Reported results for each simulation study ($n = 200$ and $n = 500$) are based on 150 Monte Carlo

runs. We ran three parallel chains for each simulated data set with 5000 iterations per chain, with the first 500 iterations discarded as burn-in period. We set thinning to keep 1500 posterior samples in each chain, thus using a total of 4500 samples for estimation and inference. In order to determine the priors for the hyperparameters, following the suggestions by[25] and[28], we examined the estimated results under several different prior choices. The priors of the hyperparameters for subject- and facility-level varying-coefficient functions in both submodels were set to $a_{1\tau_*} = a_{1\kappa_*} = 1$ and $a_{2\tau_*} = a_{2\kappa_*} = 0.005$. For the time-varying coefficient $\alpha(t)$ and baseline hazard function $h_0(t)$, $\left(a_{1\kappa_\alpha}, a_{2\kappa_\alpha}\right) = (0.1, 0.5)$ and $\left(a_{1\kappa_{h_0}}, a_{2\kappa_{h_0}}\right) = (0.1, 0.005)$, respectively, were chosen as the priors.

We evaluate the performance of the proposed procedure in estimating the time-invariant ($\sigma_S^2$ and $\sigma_F^2$) and time-varying ($\alpha$(t), $h_0(t)$, and varying-coefficient functions from the longitudinal or survival submodels, $\beta_*(t)$ or $\gamma_*(t)$, respectively) coefficients using the mean squared error (MSE) and root average squared error (RASE), respectively. RASE[24,32,33], a commonly used measure to assess the accuracy of varying-coefficient estimations, is defined as

$$RASE_{\hat{f}} = \left[\frac{1}{K}\sum_{k=1}^{K}\left\{\frac{f(t_k) - \hat{f}(t_k)}{\text{range } f(\cdot)}\right\}^2\right]^{1/2}, \quad (5)$$

where $K = 20$ and $f(\cdot)$ denotes a single varying-coefficient function. The estimated time-varying coefficient functions in longitudinal and survival submodels and the estimated baseline hazard function are depicted in Figures 1–2 along with their simultaneous credible bands from the simulation runs with the median RASE based on $n = 200$. We observe that our procedure performed well in simulation. In particular, the estimates (dashed) are close to the true functions (solid) that are mostly captured well within the simultaneous (dotted) and pointwise (dashed-dotted) credible intervals for all varying-coefficient functions. As expected, results improved (e.g., smaller bias and narrower pointwise and simultaneous credible bands) for the $n = 500$ facilities case (Figure 3). For subject-level random effect variance $\sigma_S^2$, we obtain the estimated values and the 95% credible intervals for $n = 200$ and $n = 500$ as 1.285 (1.245, 1.326) and 1.294 (1.261, 1.327) respectively. Similarly, for the facility-level random effect variance $\sigma_F^2$, the estimated values and the 95% credible intervals are 0.227 (0.182, 0.273) and 0.218 (0.181, 0.254) for $n = 200$ and $n = 500$, respectively. The results indicate that our method targets the true values closely, and similar to results on varying-coefficient functions, the bias gets smaller as the number of facilities increases. The (25th, 50th, 75th) percentiles of the RASE values for the varying-coefficient functions and the MSE values for the subject- and facility-level random effect variances are displayed in Table 1. As expected, the values for these error measures get smaller as the number of facilities increases. We also assess the coverage probabilities. Due to the clear boundary effects, that is, under-coverage at the boundaries, we present average coverage probabilities based on both simultaneous and pointwise credible bands for the time interval (0.2, 0.8) in Table 1. The coverage probabilities in this interval, which accounts for the boundary effects, are above the nominal level. These results are consistent with previous literature[34,35]

and indicates that the Bayesian credible intervals are more conservative. In addition, we observe that the coverage probabilities get larger as the number of facilities increase and the pointwise credible bands are narrower than the simultaneous credible intervals; therefore, the coverage probabilities based on those are smaller.

# 4 | APPLICATION: HOSPITALIZATION AND SURVIVAL IN DIALYSIS PATIENTS

## 4.1 | USRDS study cohort and patient characteristics

We applied the proposed BMT-JM to data from the USRDS, a national database that collects data on nearly all U.S. patients with ESKD on dialysis. The study cohort included patients of age 18 years or older, initiating dialysis between January 1, 2006 and December 31, 2008. The maximum follow-up period was 5 years, with the last follow-up date as December 31, 2013, where follow-up was truncated if a patient switched dialysis facilities. The inclusion criteria were: (1) patients who survived the first 90 days, did not have recovery of kidney function, and did not have a kidney transplant, and (2) patients who were covered by Medicare as their primary payer on Day 91 of the dialysis. Therefore, the first day of study follow-up started on Day 91, per the recommendation stated in the USRDS Researcher's Guide "90-day rule" to allow for completion of the Medicare eligibility application process and establishment of stable dialysis treatment modality[36]. The final study cohort included 292,672 observations over time on 34,030 patients in 520 dialysis facilities, where the number of patients per facility ranged from 50 to 162 (median 61, Q1–Q3 [first-third quartile]: 54–71). The overall hospitalization rate was 27.4% and the censoring was 61.4%.

For our joint modeling, the time index $t$ is the time starting from when a patient transition to dialysis. This choice is of prime interest in the dialysis population since research over the past decade have shown that patient outcomes, such as hospitalization and death, are elevated particularly in the first 1–1.5 years after patients transition to dialysis and this trajectory changes thereafter. Thus, our proposed modeling also focuses on this main time index of interest, specifically, time since initiation of dialysis. Additionally, baseline age, i.e., the age at which patients transition to dialysis has been shown to be an important risk factor for both hospitalization and death; hence, this effect modeled as baseline age is also important in this population. Furthermore, modeling baseline age at transition to dialysis (instead of a time-dependent version) is also needed to avoid potential overlap between a time-dependent age covariate and follow-up across the main time index of interest. Hence, our current model is able to capture time-varying trends of covariates on outcomes as patients stay longer on dialysis while adjusting for the effect of age at initiation of dialysis.

The mean age of the patients in the study cohort was 65 years (SD 15), and 45% of the patients were female. Common baseline comorbidities included chronic obstructive pulmonary disease (COPD; 18.7%), septicemia (10.2%), other infectious diseases (23.1%), cardiorespiratory failure (12%), coagulopathy (7.9%), and psychiatric conditions (11.2%). Among the 520 facilities, the median length of patient follow-up was 24.3 months (Q1–Q3: 21.1–27.4), and the mean number of hospitalizations was 1.8 per person-year (SD 2.2). The median marginal (unadjusted) survival was calculated as 46.5 months (3.9 years), and

we observed that the survival probability decreases noticeably for ESKD patients with additional baseline comorbidities such as COPD and/or septicemia. In terms of facility-level data, on average, the ratio of nurse-to-patient and patient care technician (PCT)-to-patient were 7.6% (SD 3.2) and 9.4% (SD 2.9), respectively.

The proposed BMT-JM was employed to study the time-varying effects of patient-level covariates on both longitudinal and survival outcomes. Patient-level covariates included age (centered), sex, and baseline comorbidities (COPD, coagulopathy, cardiorespiratory failure, septicemia, other infectious diseases/pneumonias, and psychiatric disorders). The facility-level covariates included in both submodels were nurse-to-patient ratio and PCT-to-patient ratio (both centered). Following from (3), the survival submodel also had the hospitalization risk score as a covariate.

### 4.2 | Results

The estimated time-varying coefficient and baseline hazard functions were obtained using the Bayesian P-splines approach with 20 equally spaced knots and a second-order penalty. To obtain the posterior samples, we ran three parallel chains with 5000 iterations per chain, where 500 iterations were discarded as burn-in period. The thinning was chosen to keep 1500 posterior samples in each chain; hence we used 4500 samples for estimation and inference. The priors of the hyperparameters for subject- and facility-level varying-coefficient functions in the longitudinal submodel were set to $\left(a_{1\tau_*}, a_{2\tau_*}\right) = (1, 0.005)$. In the survival submodel, we used $\left(a_{1\kappa_{X_\omega}}, a_{2\kappa_{X_\omega}}\right) = (1, 0.0005)$ and $\left(a_{1\kappa_{Z_\upsilon}}, a_{2\kappa_{Z_\upsilon}}\right) = (10, 0.0005)$ for the subject- and facility-level varying-coefficient functions, respectively. For the baseline hazard function and $\alpha(t)$, we selected $\left(a_{1\kappa_\alpha}, a_{2\kappa_\alpha}\right) = (1, 0.005)$ and $\left(a_{1\kappa_{h_0}}, a_{2\kappa_{h_0}}\right) = (10, 0.0005)$, respectively, as the priors. Lastly, the priors for the variance components were set to $(a_{1*}, a_{2*}) = (10, 0.0005)$. The subject- and facility-level random effect variances were estimated as $\sigma_S^2 = 1.344$ and $\sigma_F^2 = 0.106$ with standard errors 0.027 and 0.015, respectively.

The estimated time-varying effects of patient-level risk factors (solid line) on longitudinal hospitalizations are displayed in Figure 4, along with 95% simultaneous credible bands (dashed lines) and pointwise credible intervals (dotted lines). The estimated intercept $\hat{\beta}_{X_1}(t)$ in the longitudinal submodel is increasing over time (Figure 4(a)). More specifically, the odds of hospitalization $\left(\exp\left\{\hat{\beta}_{X_1}(t)\right\}\right)$ for male patients initiating dialysis at mean age 65 with no comorbidities and treated at a typical dialysis facility (with average nurse-to-patient and PCT-to-patient ratio of 7.6% and 9.4%, respectively) is increasing over time, post-dialysis transition. Females have a higher odds of hospitalization compared to males, but this declines over the five year follow-up with highest time-varying odds ratio $OR(t) = \exp\left\{\hat{\beta}_2(t)\right\} \sim 1.23$ during the first year on dialysis (Figure 4(b)). Older age at transition to dialysis is associated with a higher odds of hospitalization starting at about 12 months (i.e., after the fragile first year transition to dialysis time period; Figure 4(c)).

As expected, all comorbidities are associated with significantly higher odds of hospitalization (Figure 4(d)–(i)). In particular, the effects of chronic comorbid conditions,

such as COPD and psychiatric conditions, are relatively stable over time and are associated with significantly higher odds hospitalization (e.g., $OR(t)$ ranges from 1.47 to 1.69 for COPD and 1.23 to 1.47 for psychiatric conditions). The effects of acute comorbidities, such as septicemia and other infectious diseases/pneumonia, decrease as patients remain longer on dialysis, but the highest risk period is during the first 12 months after transitioning to dialysis: $OR(t)$ : 1.37 − 1.91 for septicemia and 1.38 − 1.72 for other infections, $t < 12$ months. Of note, although the risk of hospitalization sharply declines after the first year, it remains significantly higher compared to those without prior acute infections ($OR(t)$ : 1.21 − 1.37 for $t > 12$). Similarly, coagulopathy and cardiorespiratory failure have significant effects on hospitalization during the first year post-dialysis transition.

The effects of patient-level covariates on the hazard of death are displayed in Figure 5. We observe that nearly all comorbidities (except for coagulopathy) have significant effects on the risk of death during the first year of dialysis, with the highest hazard of death observed in the several months immediately following transition to dialysis. The effects of comorbidities decrease over time as patients stay longer on dialysis. For example, COPD has a significant effect on survival approximately until the end of the third year of dialysis, with estimated time-varying hazard ratio, $HR(t) = \exp\{\hat{\gamma}_*(t)\}$, varying from 1.16 − 1.27 for $t < 30$ months. The effect trajectories of cardiorespiratory failure and septicemia on survival remain significant during the first two years after dialysis transition. Also, as expected, older age at dialysis transition is associated with an increased hazard of death. The estimated time-varying effect of hospitalization risk score, $\hat{\alpha}(t)$, on the hazard of death is presented in Figure 5(i). Although the simultaneous credible bands show a non-significant effect (except for months 40–50), the pointwise credible intervals demonstrate a significant effect on survival up to about 20 months and the highest point estimate of $HR(t) \sim 1.25$ at about 10 months post-dialysis. We note that the estimated baseline hazard of death was approximately constant over time (not shown).

Finally, at the facility level, the estimated time-varying effects of nurse-to-patient ratio and PCT-to-patient ratio on longitudinal and survival outcomes are presented in Figure 6. The nurse-to-patient ratio is not found to be significantly associated with either outcome (Figure 6 (a), (c)). The PCT-to-patient ratio significantly affects the hazard of death (Figure 6 (d)) and hospitalization (Figure 6 (b)) approximately until 20 and 30 months of dialysis, respectively, where a higher ratio of technicians-to-patients is associated with a lower hazard of death and odds of hospitalization. However, these effect sizes are small relative to the effect sizes of patient-level risk factors, similar to prior findings [2,3].

## 5 | DISCUSSION

Motivated by the need to model three-level hierarchical data from the USRDS of the dialysis population, we developed a Bayesian multilevel time-varying framework to jointly model patient longitudinal hospitalization and survival. This work fills several critical methodological gaps in the joint modeling literature. First, it provides a joint modeling approach that (a) accommodates higher-level hierarchical data, especially for large data (providing a computationally feasible approach to handling ultra high-dimensional REs) and (b) allows for multilevel risk factors at the individual-level and the cluster-level (dialysis

facility). Second, to our knowledge, this is the first joint model of longitudinal and survival outcomes with time-varying coefficients for higher-level hierarchical data. The need to develop such a joint model was illustrated by our need to investigate the potential time-varying effects of both patient-level and dialysis facility-level risk factors on hospitalization and survival in the dialysis population. Indeed, as detailed in the Results section, the analysis identified risk factors as well as specific time periods of elevated risk of hospitalization and death. For example, a history of acute comorbid conditions (infections, septicemia) prior to transition to dialysis is associated with a substantial risk of hospitalization, especially in the first year, suggesting a more aggressive monitoring of patients with such profile to reduce the likelihood of hospitalization during this fragile time period. On the other hand, chronic comorbid conditions such as COPD and psychiatric conditions have sustained and large effects on hospitalization throughout the time periods after patients transition to dialysis. Thus, these findings suggest a tailored approach to dialysis patient treatment and monitoring based on patient characteristics (e.g., type of comorbid conditions) and specific time periods of elevated hospitalization and mortality risk after transition to dialysis.

We note that in our joint modeling approach, the model for the hazard of mortality includes the longitudinal process (history of hospitalization) by choice. The reason for this choice is motivated by the fact that for the dialysis population, hospitalization and survival is intricately related and frequent hospitalizations in this population (at about twice per year) is a major source of both morbidity and mortality which have been well documented in our own works and others [5,6,4,1]. Therefore, in this population, understanding time-varying effects of risk factors as well as hospitalization risk is key to developing effective patient monitoring strategies to reduce *mortality*; thus, hospitalization as a time-varying risk factor is itself an important recognized comorbid condition of equal importance relative to other traditional chronic and acute comorbid conditions in this population. The results displayed in Figure 5 show that the magnitude of the effect of hospitalization risk on mortality is roughly commensurate with the effect sizes of major chronic comorbid conditions such as COPD and cardiorespiratory failure during the first 18 months after transition to dialysis. Generally, the inclusion of the longitudinal process into the hazard model should be motivated by the specific data application. Other examples of this modeling choice include works by[37] and[20]. Finally, although not of primary interest in modeling survival in the dialysis population, we note that the effect of traditional risk factors, such as COPD and cardiorespiratory failure, absent the longitudinal process, may be of interest in other contexts (e.g., other populations). In such a case, exclusion of the longitudinal process in the hazard model would be an appropriate choice.

## ACKNOWLEDGMENTS

This study was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK092232 D. S., D. V. N., E. K., S. B., C. M. R., and Y. L.). The data reported here have been supplied by the United States Renal Data System (USRDS). The interpretation and reporting of the data presented here are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the U.S. government.

## DATA AVAILABILITY STATEMENT

The release of the data used in this paper is governed by the National Institute of Diabetes and Digestive and Kidney Diseases through the USRDS Coordinating Center. The data can be requested from the USRDS through a data use agreement.

## References


1. United States Renal Data System. USRDS2021 annual data report: epidemiology of kidney disease in the United States. Technical report Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2021.
2. Chen Y, Rhee C, Şentürk D. et al. Association of US dialysis facility staffing with profiling of hospital-wide 30-day unplanned readmission. Kidney Dis 2019;5(3):153–162.
3. Li Y, Nguyen DV, Kürüm E, et al. A multilevel mixed effects varying coefficient model with multilevel predictors and random effects for modelling hospitalization risk in patients on dialysis. Biometrics 2020;76(3):924–938. [PubMed: 31856300]
4. Li Y, Nguyen DV, Chen Y, Rhee CM, Kalantar-Zadeh K,Şentürk D. Modeling time-varying effects of multilevel risk factors of hospitalizations in patients of dialysis. Stat Med 2018;37(30):4707–4720. [PubMed: 30252153]
5. Ester JP, Nguyen DV, Dalrymple LS, Mu Y, Şentürk D. Time-varying effect modelling with longitudinal data truncated by death: conditional models, interpretations, and interpretations, and inference. Stat Med 2016;11(35):1834–1847.
6. Estes JP, Nguyen DV, Chen Y, et al. Time-dynamic profiling with application to hospital readmission among patients on dialysis. Biometrics 2018;4(74):1383–1394.
7. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. Stat Sin 2004;14:809–834.
8. Rizopoulos D Joint Models for Longitudinal and Time-to-Event Data: With Applications in R Boca Raton, FL: CRC Press; 2012.
9. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics 1997;53:330–339. [PubMed: 9147598]
10. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. Biostatistics 2000;1(4):465–480. [PubMed: 12933568]
11. Tsiatis AA, Davidian M. A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. Biometrika 2001;88(2):447–458.
12. Song X, Davidian M, Tsiatis AA. A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. Biometrics 2002;58(4):742–753. [PubMed: 12495128]
13. Hsieh F, Tseng Y, Wang J. Joint modeling of survival and longitudinal data: likelihood approach revisited. Biometrics 2006;62(4):1037–1043. [PubMed: 17156277]
14. Rizopoulos D, Verbeke G, Molenberghs G. Shared parameter models under random effects misspecification. Biometrika 2008;95(1):63–74.
15. Mauff k Steyerberg E, kardys I, Boersma E, Rizopoulos D Joint models with multiple longitudinal outcomes and a time-to-event outcome: a corrected two-stage approach. Stat Comput 2020;30(4):999–1014.
16. Liu L, Ma JZ, O'Quigley J. Joint analysis of multi-level repeated measures data and survival: an application to the end stage renal disease (ESRD) data. Stat Med 2008;27(27):5679–5691. [PubMed: 18693300]
17. Brilleman SL, Crowther MJ, Moreno-Betancur M, et al. Joint longitudinal and time-to-event models for multilevel hierarchical data. Stat Methods Med Res 2019;12(28):3502–3515.
18. kürüm E, Nguyen DV, Li Y, Rhee CM, kalantar-Zadeh k, Şentürk D Multilevel joint modeling of hospitalization and survival in patients on dialysis. Stat 2021;10(1):e356.
19. Song X, Wang C. Semiparametric approaches for joint modeling of longitudinal and survival data with time–Varying coefficients. Biometrics 2008;64(2):557–566. [PubMed: 17725812]

20. Andrinopoulou E, Eilers PH, Takkenberg JJ, Rizopoulos D. Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. Biometrics 2018;74(2):685–693. [PubMed: 29092100]
21. Piulachs X, Andrinopoulou E, Guillén M, Rizopoulos D. A Bayesian joint model for zero-inflated integers and left-truncated event times with a time-varying association: applications to senior health care. Stat Med 2021;40(1):147–166. [PubMed: 33104241]
22. Cleveland W, Grosse E, Shyu W. Local Regression Models Pacific Grove, CA: Wadsworth & Brooks/Cole; 1992.
23. Hastie T, Tibshirani R. Varying-coefficient models. J Royal Stat Soc Ser B 1993;55(4):757–796.
24. Cai Z, Fan J, Li R. Efficient estimation and inferences for varying-coefficient models. J Am Stat Assoc 2000;95(451):888–902.
25. Lang S, Brezger A. Bayesian P-splines. J Comput Graph Stat 2004;13(1):183–212.
26. Eilers PH, Marx BD. Flexible smoothing with B-splines and penalties. Stat Sci 1996;11(2):89–121.
27. Jullion A, Lambert P. Robust specification of the roughness penalty prior distribution in spatially adaptive Bayesian P-splines models. Comput Stat Data Anal 2007;51(5):2542–2558.
28. Brezger A, Lang S. Generalized structured additive regression based on Bayesian P-splines. Comput Stat Data Anal 2006;50(4):967–991.
29. Crainiceanu CM, Ruppert D, Carroll RJ, Joshi A, Goodner B. Spatially adaptive Bayesian penalized splines with heteroscedastic errors. J Comput Graph Stat 2007;16(2):265–288.
30. Plummer M, Stukalov A, Denwood M. rjags: Bayesian graphical models using MCMC. R package version, 4.10; 2019.
31. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. Stat Med 2005;24(11):1713–1723. [PubMed: 15724232]
32. kürüm E, Li R, Wang Y, Ş entürk D. Nonlinear varying-coefficient models with applications to a photosynthesis study. J Agric Biol Environ Stat 2014;19(1):57–81. [PubMed: 24976756]
33. kürüm E, Li R, Shiffman S, Yao W. Time-varying coefficient models for joint modeling binary and continuous outcomes in longitudinal data. Stat Sin 2016;26(3):979. [PubMed: 27667908]
34. Cox DD. An analysis of Bayesian inference for nonparametric regression. Ann Stat 1993;21(2):903–923.
35. krivobokova T, kneib T, Claeskens G. Simultaneous confidence bands for penalized spline estimators. J Am Stat Assoc 2010;105(490):852–863.
36. United States Renal Data System. USRDS 2014 annual data report: epidemiology of kidney disease in the United States. Technical report, Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and kidney Diseases.
37. Rizopoulos D, Taylor JM, Van Rosmalen J, Steyerberg EW, Takkenberg JJ. Personalized screening intervals for biomarkers using joint models for longitudinal and survival data. Biostatistics 2016;17(1):149–164. [PubMed: 26319700]

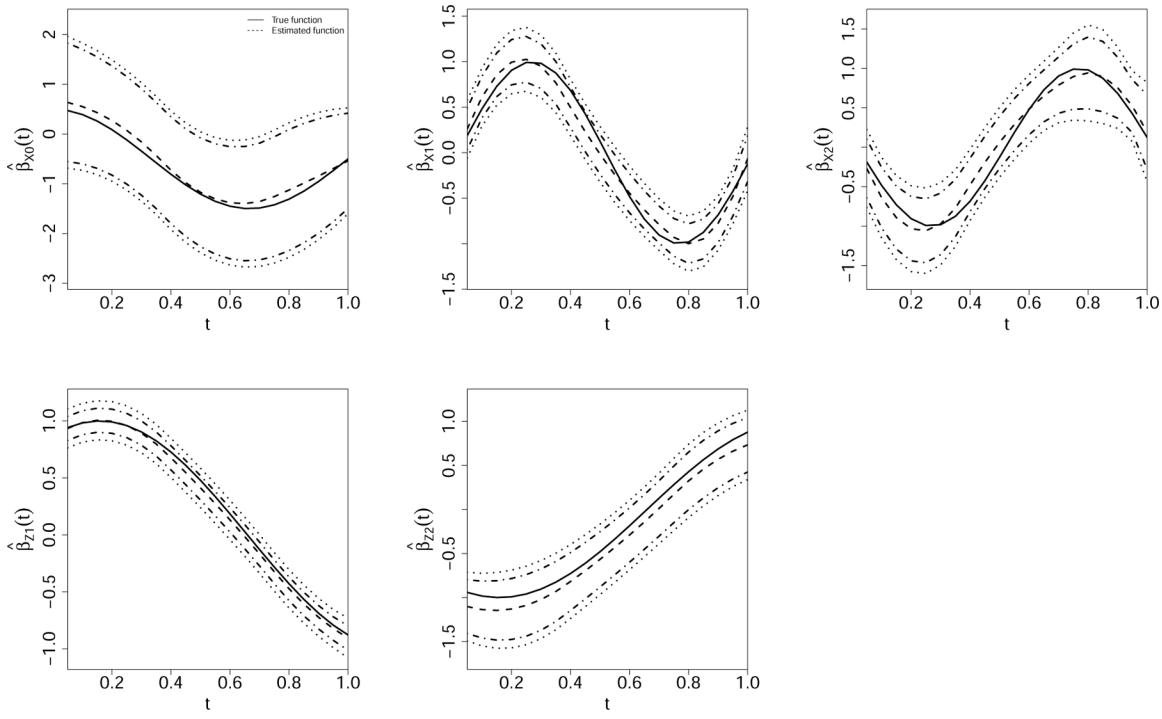


**FIGURE 1.**

Estimated time-varying coefficient functions (dashed) in the longitudinal submodel from the simulation runs with the median RASE among 150 Monte Carlo runs for $n = 200$ facilities overlaying the true functions (solid) along with 95% simultaneous (dotted) and pointwise (dashed-dotted) credible intervals.

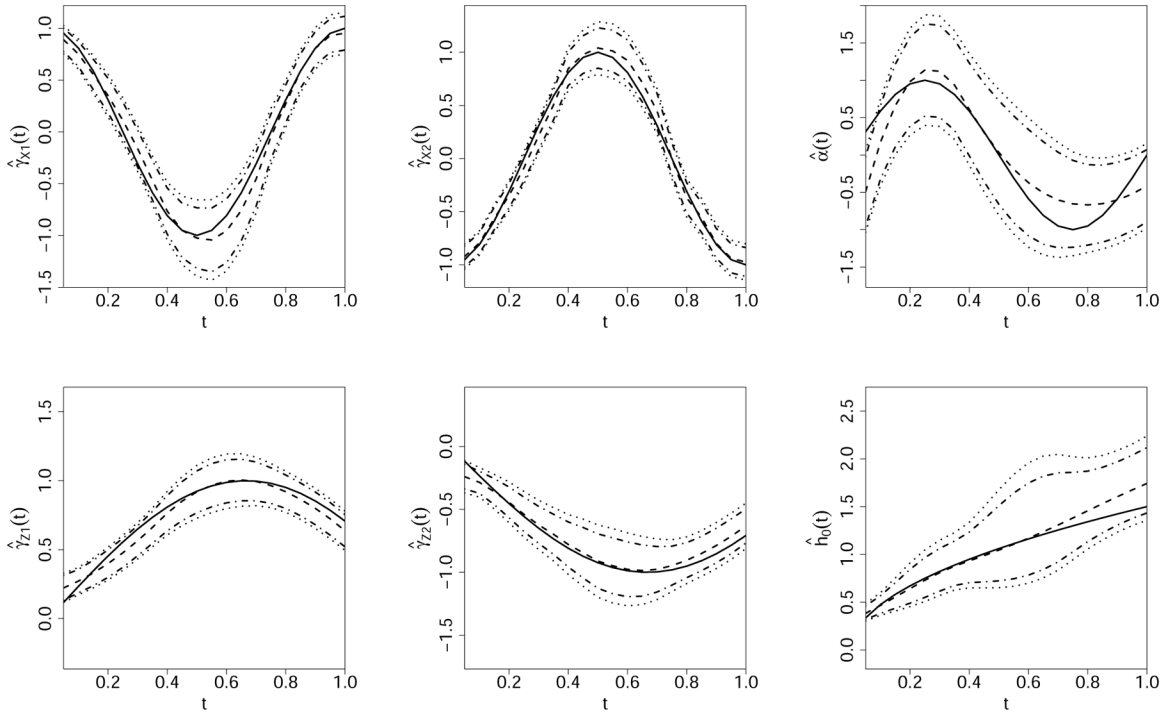


**FIGURE 2.**

Estimated time-varying coefficient functions and baseline hazard function (dashed) in the survival submodel from the simulation runs with the median RASE among 150 Monte Carlo runs for $n = 200$ facilities overlaying the true functions (solid) along with 95% simultaneous (dotted) and pointwise (dashed-dotted) credible intervals.

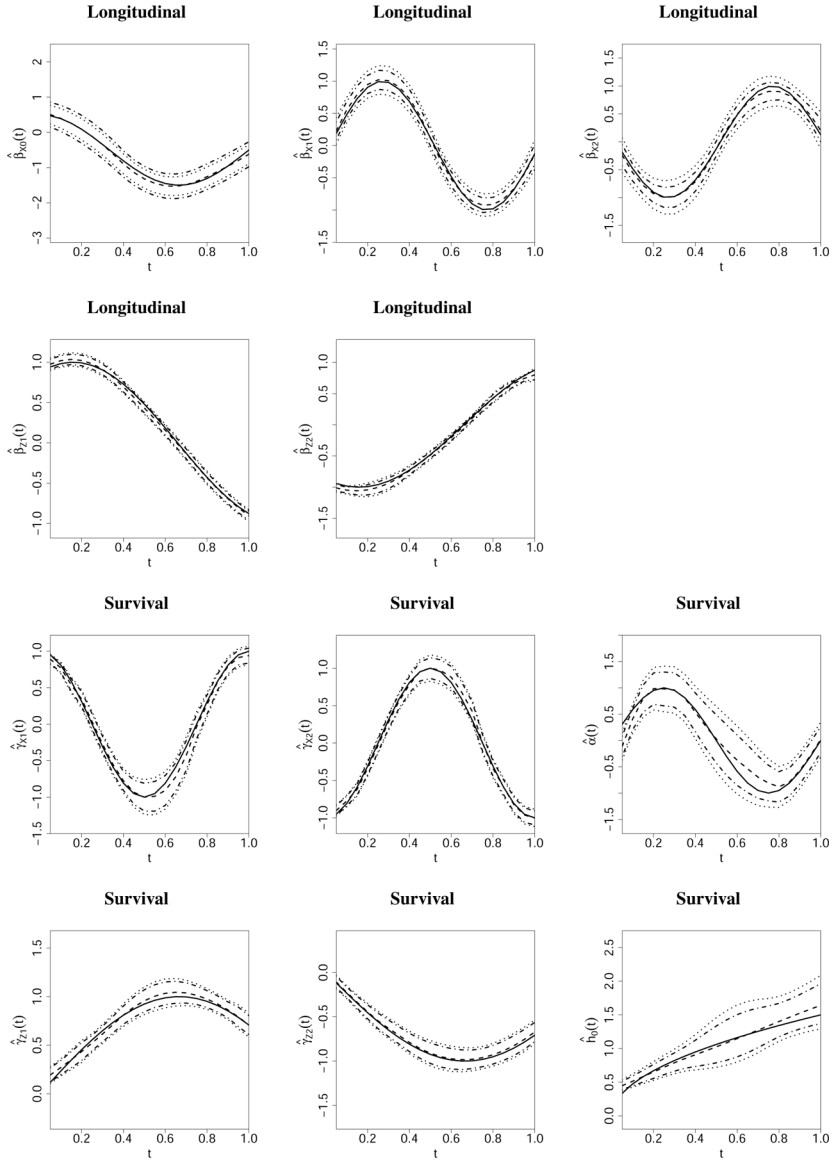


**FIGURE 3.**

Estimated time-varying functions (dashed) in the longitudinal and survival submodels from the simulation runs with the median RASE among 150 Monte Carlo runs for $n = 500$ facilities overlaying the true functions (solid) along with 95% simultaneous (dotted) and pointwise (dashed-dotted) credible intervals.

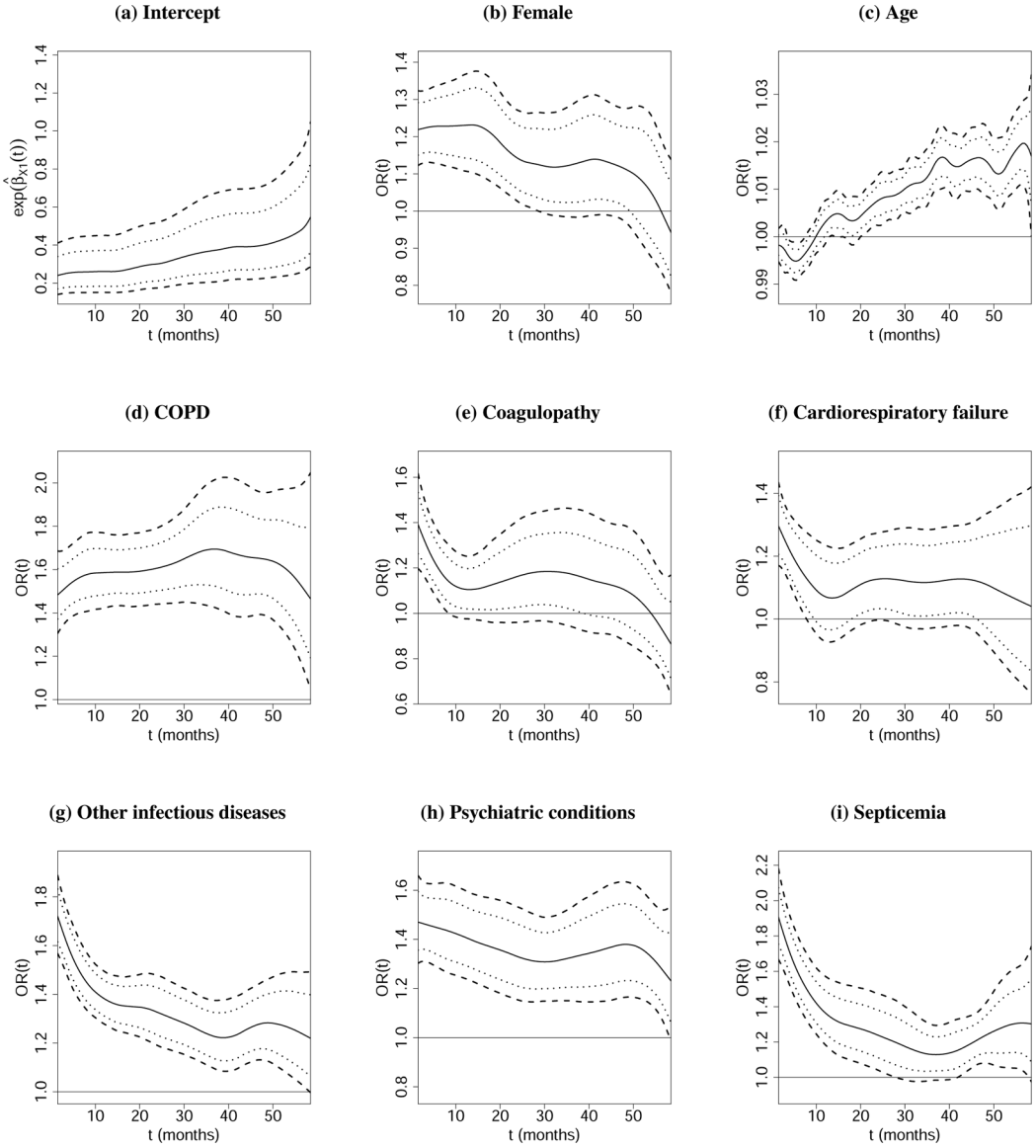


**FIGURE 4.**

Estimated patient-level effects on hospitalization, time-varying odds ratio $OR(t) = \exp\left\{\hat{\beta}_{X_*}(t)\right\}$, (solid) along with their 95% simultaneous (dashed) and pointwise (dotted) credible bands.

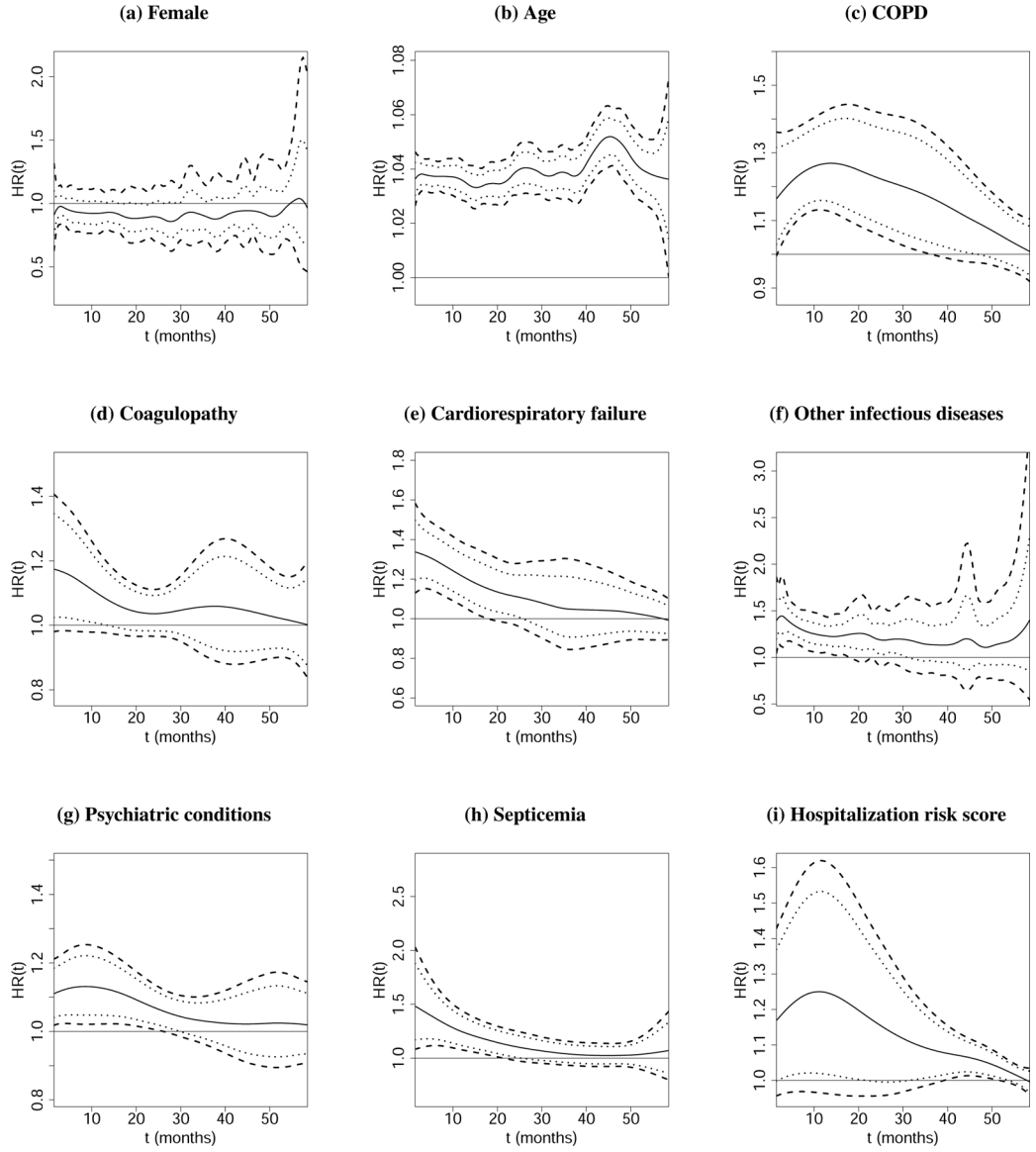


**FIGURE 5.**

Estimated patient-level effects on survival, time-varying hazard ratios $HR(t) = \exp\{\hat{\gamma}_{X_*}(t)\}$ and $HR(t) = \exp\{\hat{\alpha}(t)\}$, (solid) along with their 95% simultaneous (dashed) and pointwise (dotted) credible bands.

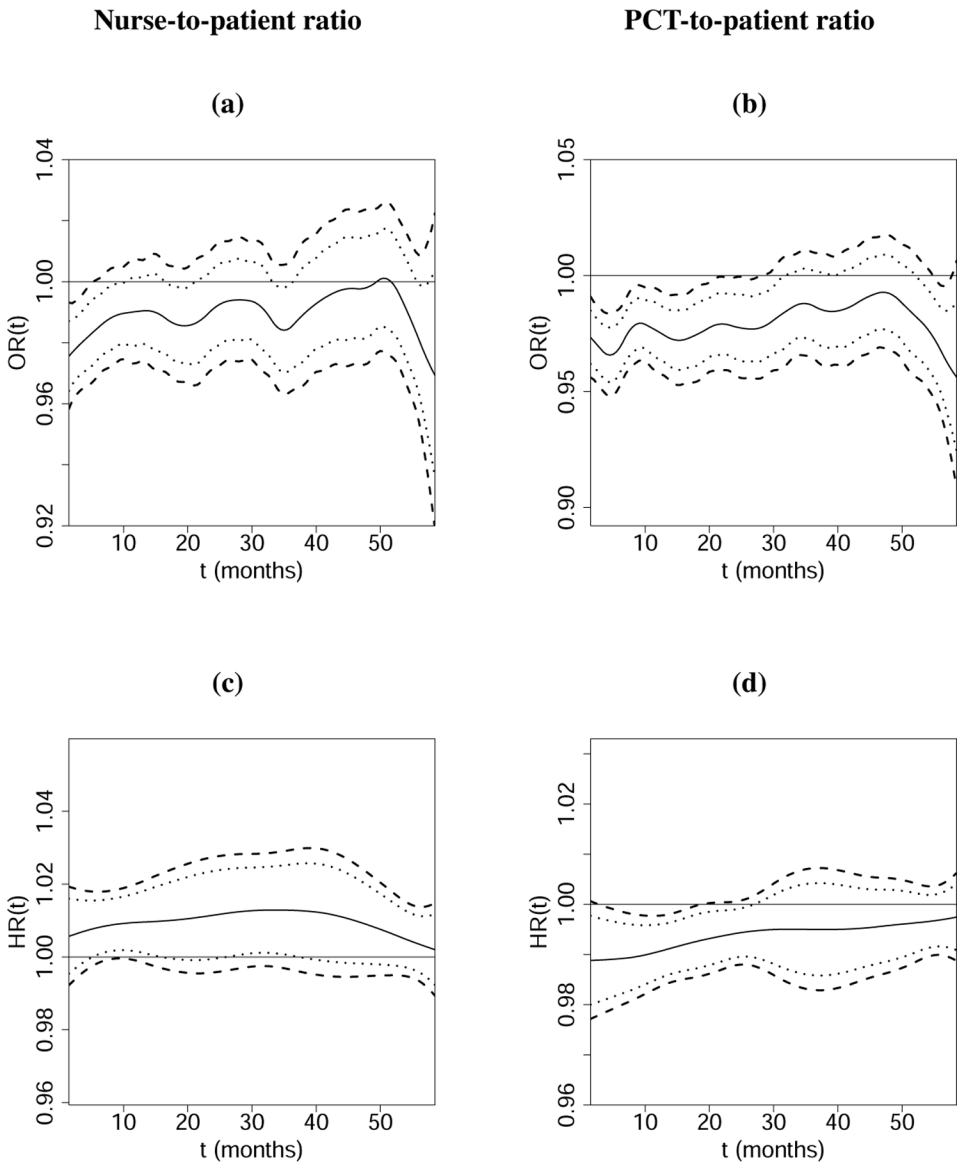


**FIGURE 6.**

Estimated facility-level effects on (a, b) hospitalization, time-varying odds ratio $OR(t) = \exp\{\hat{\beta}_{Z_*}(t)\}$, and (c, d) on survival, time-varying hazard ratio $HR(t) = \exp\{\hat{\gamma}_{Z_*}(t)\}$, (solid) along with their 95% simultaneous (dashed) and pointwise (dotted) credible bands.

## TABLE 1

The (25th, 50th, 75th) percentiles of the root average squared error (RASE) and mean squared error (MSE) for the time-varying and time-invariant estimates, respectively, along with average coverage probabilities of the 95% simultaneous ($CP_S$) and pointwise (CP) credible band (computed for the time interval (0.2, 0.8) for time-varying coefficients) based on 150 Monte Carlo runs.

| | Number of facilities, $n$ = 200 | | | Number of facilities, $n$ = 500 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **RASE/MSE** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **$CP_S$(%)** | **CP(%)** | **$CP_S$(%)** | **CP(%)** |
| Longitudinal | | | | | | | | | | |
| $\beta_{X_0}(t)$ | 0.043 | 0.067 | 0.104 | 0.016 | 0.031 | 0.078 | 97.3 | 96.6 | 97.5 | 97.0 |
| $\beta_{X_1}(t)$ | 0.049 | 0.051 | 0.053 | 0.021 | 0.022 | 0.023 | 97.2 | 96.3 | 97.3 | 96.5 |
| $\beta_{X_2}(t)$ | 0.056 | 0.059 | 0.062 | 0.022 | 0.024 | 0.028 | 98.1 | 97.0 | 98.2 | 97.2 |
| $\beta_{Z_1}(t)$ | 0.015 | 0.021 | 0.028 | 0.011 | 0.014 | 0.016 | 97.5 | 95.6 | 97.7 | 96.0 |
| $\beta_{Z_2}(t)$ | 0.027 | 0.064 | 0.099 | 0.010 | 0.021 | 0.032 | 97.0 | 95.2 | 97.1 | 95.8 |
| Survival | | | | | | | | | | |
| $\gamma_{X_1}(t)$ | 0.021 | 0.043 | 0.112 | 0.024 | 0.029 | 0.041 | 95.3 | 95.1 | 96.1 | 95.5 |
| $\gamma_{X_2}(t)$ | 0.023 | 0.033 | 0.123 | 0.012 | 0.022 | 0.033 | 95.6 | 95.3 | 96.0 | 95.5 |
| $\gamma_{Z_1}(t)$ | 0.034 | 0.057 | 0.095 | 0.018 | 0.037 | 0.046 | 95.5 | 95.2 | 96.0 | 95.7 |
| $\gamma_{Z_2}(t)$ | 0.038 | 0.053 | 0.095 | 0.008 | 0.022 | 0.049 | 98.2 | 96.3 | 98.2 | 97.0 |
| $\alpha(t)$ | 0.111 | 0.142 | 0.185 | 0.052 | 0.076 | 0.120 | 96.9 | 95.7 | 97.4 | 96.8 |
| $h_0(t)$ | 0.034 | 0.083 | 0.150 | 0.032 | 0.055 | 0.062 | 97.1 | 96.3 | 97.2 | 96.5 |
| Variance components | | | | | | | | | | |
| $\sigma_S^2$ | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | – | 95.0 | – | 95.1 |
| $\sigma_F^2$ | <.001 | 0.002 | 0.007 | <.001 | 0.001 | 0.005 | – | 91.3 | – | 93.4 |