Title

Modern Modeling Techniques for Adverse Events in Chronic Kidney Disease Patients

Author

Qian, Qi

Publication Date

2025

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Modern Modeling Techniques for Adverse Events in Chronic Kidney Disease Patients

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Biostatistics

by

Qi Qian

2025

ABSTRACT OF THE DISSERTATION

Modern Modeling Techniques for Adverse Events in
Chronic Kidney Disease Patients

by

Qi Qian
Doctor of Philosophy in Biostatistics
University of California, Los Angeles, 2025
Professor Damla Şentürk, Chair

Chronic kidney disease (CKD) affects nearly 15% (37 million) of adults in the United States. End-stage kidney disease (ESKD), the final stage of CKD, impacted over 807,920 individuals in 2020, with approximately 70% receiving dialysis – a life-sustaining treatment. Among CKD patients, the progressive decline in kidney function is closely linked to hospitalizations due to cardiovascular (CV) disease and terminal events such as kidney failure and death. Therefore, jointly modeling the correlated multivariate outcomes of hospitalization and mortality, along with their spatiotemporal patterns and associated risk factors, is essential for improving targeted patient monitoring and intervention.

In the first chapter, using national data from the United States Renal Data System (USRDS), we propose a novel multivariate spatiotemporal functional principal component analysis (MST-FPCA) model to study joint spatiotemporal patterns in hospitalization and mortality rates among dialysis patients. This model leverages a multivariate Karhunen-Loéve expansion to capture leading directions of variation across time while inducing spatial correlation among region-specific scores. We develop an efficient estimation procedure that combines only univariate principal component decompositions with a Markov chain Monte Carlo (MCMC) framework to target spatial dependencies. The finite sample performance of the proposed method is studied through simulation studies. Application to the USRDS data identifies geographic "hot spots" in the U.S. with elevated hospitalization and/or mortality

rates and highlights time periods of heightened clinical risk.

Building on these findings, the second chapter explores the time-varying effects of regional-level covariates on hospitalization and mortality. We propose a multivariate varying coefficient spatiotemporal model (MV-VCSTM) to study the dynamic effects of risk factors (e.g., urbanicity, area deprivation index) on hospitalization and mortality rates as functions of time on dialysis. The proposed model flexibly accommodates time-varying covariate effects on the mean structure while modeling the spatiotemporal correlations of the residuals for efficient inference. Estimation is achieved through a fusion of functional principal component analysis and MCMC techniques, following basis expansions for varying coefficient functions and a multivariate Karhunen-Loéve expansion for region-specific random deviations. The finite sample performance of the proposed method is studied through extensive simulations. Application to the USRDS data uncovers significant regional risk factors and identifies critical time periods on dialysis and geographic areas with increased hospitalization and mortality risks.

In the third chapter, we shift focus to individual-level disease progression. We introduce a novel Bayesian multivariate joint model (BM-JM) for the correlated outcomes of longitudinal kidney function (measured by estimated glomerular filtration rate, eGFR), recurrent CV events, and competing-risk terminal events (ESKD and death). The proposed modeling framework facilitates joint estimation of risk factors for each outcome and enables dynamic, personalized prediction of cumulative incidence probabilities for competing risks. These predictions are informed by a subject's baseline characteristics and their history of longitudinal eGFR and recurrent CV events. Estimation and prediction are conducted within a flexible Bayesian framework using MCMC methods. We assess predictive performance via the dynamic area under the receiver operating characteristic curve (AUC) and expected Brier score (BS). Simulation studies demonstrate the model's robustness. Application to data from the Chronic Renal Insufficiency Cohort (CRIC) study illustrates its practical utility and further emphasizes the value of integrating both longitudinal and recurrent event data to improve prognostic accuracy in CKD populations.

The dissertation of Qi Qian is approved.

Jin Zhou

Donatello Telesca

Danh V. Nguyen

Damla Şentürk, Committee Chair

University of California, Los Angeles

2025

This dissertation is dedicated to my families
for their endless love and unconditional support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Over the past six years in the Department of Biostatistics at the University of California, Los Angeles (UCLA), I have been incredibly fortunate to receive unwavering support, encouragement, and mentorship from many individuals. This dissertation would not have been possible without their invaluable contributions, and I am profoundly grateful to each of them. I also gratefully acknowledge the award from the UCLA Graduate Division and a research grant from the National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK092232), which made my doctoral work possible.

First and foremost, I would like to express my deepest gratitude to my advisor and committee chair, Dr. Damla Şentürk. She has been an extraordinary mentor – steadfast, generous, and inspiring. Her guidance has shaped not only my academic journey but also my personal and professional growth. From teaching me how to develop statistical methodologies and conduct collaborative research, to mentoring me in scientific writing and career advancement, she has consistently led by example. Her encouragement and wisdom have empowered me to grow into an independent and confident researcher.

I am also deeply appreciative of my dissertation committee members, as well as many professors across UC campuses who have supported my academic journey. I am especially grateful to Dr. Danh V. Nguyen, whose contributions were foundational to my work. He not only provided well-prepared datasets but also offered generous and insightful guidance throughout the research process. His deep understanding of both the statistical and clinical contexts, along with his thoughtful feedback on my research goals and manuscripts, significantly shaped the direction and clarity of my work. I am thankful to Dr. Donatello Telesca, whose teaching and discussions sparked my interest in Bayesian statistics and functional data analysis – tools that became central to my dissertation. His expertise, along with his suggestions on research planning and manuscript development, greatly enhanced the methodological rigor of my work. I thank Dr. Jin Zhou for her valuable suggestions on survival analysis, as well as her constructive feedback throughout the dissertation process. Her advice helped refine key modeling decisions and improved the structure and clarity of my writing. I

am also especially thankful to Dr. Sudipto Banerjee for his thought-provoking feedback and for broadening my understanding of spatial statistics, as well as to Dr. Esra Kürüm for her patient and thoughtful guidance on joint survival modeling and parallel computing, which was critical to the development of several key components of my methodology.

I also wish to thank Roxy Naranjo, our exceptional student affairs officer, whose support has been instrumental throughout my time at UCLA. Her kindness and efficiency helped me navigate every step of the program – from coursework to dissertation – with confidence and ease. I am equally grateful to my fellow students for the stimulating discussions, collaborative spirit, and friendships that have made this journey not only intellectually rewarding but also personally fulfilling.

Finally, and most importantly, I owe my deepest thanks to my family. Their unwavering love, faith, and support have carried me through every challenge and triumph. Studying abroad and pursuing a Ph.D. has been a demanding endeavor, but their belief in me has been my greatest source of strength. This achievement is as much theirs as it is mine.

VITA

2001–2008 M.D. (Clinical Medicine), Zhejiang University, China.

2013–2017 Ph.D. (Neurology), Zhejiang University, China.

2019–2021 M.S. (Biostatistics), University of California, Los Angeles, California.

2021–2025 Graduate Student Researcher, Department of Biostatistics, University of California, Los Angeles, California.

PUBLICATIONS

Boland, J.[†], Qian, Q.[†], Telesca, D., Jeste, S.S., Dickinson, A. and Şentürk, D. (2025). Central posterior envelopes for Bayesian longitudinal functional principal component analysis. *Statistics in Biosciences*, invited revision.

Qian, Q., Nguyen, D.V., Kürüm, E., Banerjee, S., Rhee, C.M. and Şentürk, D. (2025). Bayesian multivariate joint modeling of longitudinal, recurrent,and competing risk terminal events in patients with chronickidney disease. *Journal of Data Science*, https://doi.org/10.6339/25-JDS1182.

McDonald M.N., Qian, Q., Ferrario A.C., Şentürk, D., Iyer S., and Jeste S. (2025). Developmental trajectories in high-risk NICU graduates during the first year of life. *Early Human Development*, https://doi.org/10.1016/j.earlhumdev.2024.106183.

Kürüm, E., Nguyen, D.V., Qian, Q., Banerjee, S., Rhee, C.M. and Şentürk, D. (2024). Spatiotemporal multilevel joint modeling of longitudinal and survival outcomes in end-stage

kidney disease. *Lifetime Data Analysis*, https://doi.org/10.1007/s10985-024-09635-w.

Kürüm, E., Kwan B., Qian Q., Banerjee, S., Rhee, C.M., Nguyen, D.V., and Şentürk, D. (2024). A Bayesian joint model of longitudinal kidney disease progression, recurrent cardiovascular events, and terminal event in patients with chronic kidney disease. *Statistics in Biosciences*, https://doi.org/10.1007/s12561-024-09429-6.

Qian, Q., Nguyen, D.V., Kürüm, E., Rhee, C.M., Banerjee, S., Li, Y. and Şentürk, D. (2024). Multivariate varying coefficient spatiotemporal model. *Statistics in Biosciences*, https://doi.org/10.1007/s12561-024-09419-8.

Qian, Q., Nguyen, D.V., Telesca, D., Kürüm, E., Rhee, C.M., Banerjee, S., Li, Y. and Şentürk, D. (2024). Multivariate spatiotemporal functional principal component analysis for modeling hospitalization and mortality rates in the dialysis population. *Biostatistics*, https://doi.org/10.1007/s12561-024-09419-8.

Nguyen, D.V., Qian, Q., You, A.S., Kürüm, E., Rhee, C.M. and Şentürk, D. (2023). High-dimensional fixed effects profiling models and applications in end-stage kidney disease patients: Current state and future directions. *International Journal of Satatistics in Medical Research*, https://doi.org/10.6000/1929-6029.2023.12.24.

Kwan, B., Sugar, C.A., Qian, Q., Shic, F., Naples, A., Johnson, S.P., Webb, S.J., Jeste, S.S., Faja, S., Levin, A.R., Dawson, G., McPartland J.C. and Şentürk, D. (2023). Constrained multivariate functional principal components analysis for novel outcomes. *Statistics in Biosciences*, https://doi.org/10.1007/s12561-023-09399-1.

CHAPTER 1

Multivariate Spatiotemporal Functional Principal Component Analysis for Modeling Hospitalization and Mortality Rates in the Dialysis Population

1.1 Introduction

End-stage kidney disease (ESKD) affected more than 809,103 individuals in the United States as of 2019, a 41% increase from 2009. About 70% of patients with incident ESKD were on dialysis, a life-sustaining treatment (USRDS, 2021). The mortality rate among patients receiving dialysis is higher than Medicare populations with heart failure, acute myocardial infarction or cancer. Moreover, patients on dialysis have a high burden of complex comorbid conditions and experience frequent hospitalizations (at about twice per year), which is a major source of morbidity and mortality (USRDS, 2021). Hence, hospitalization and survival are intricately related in the dialysis population.

ESKD patients typically remain on dialysis for the rest of their lives (or until kidney transplantation) and their needs change as they stay longer on dialysis. Our own works and others have shown that mortality risk and hospitalization rates change over time after patients transition to dialysis (higher hospitalization and mortality rates have been reported within the first year and first six months on dialysis, respectively (Estes et al., 2018; Li et al., 2018)). In addition to temporal variation, hospitalization and mortality rates also vary significantly across the U.S., contributing to spatial variation (Li et al., 2021). Understanding the geospatial patterns of outcomes for dialysis patients has been an important objective of the United States Renal Data System (USRDS) annual reporting. Hence, studying the

spatiotemporal trends in the multivariate outcome of hospitalization risk and mortality rates for regions across U.S. is an important goal in identifying "hot spots"/regions and critical time periods (after transitioning to dialysis) of high risk and elevated rates for more targeted patient monitoring.

There is extensive literature on multivariate spatiotemporal modeling, particularly in environmental health, criminology, road traffic analysis, and disease mapping. The existing multivariate space-time approaches can broadly be classified into four settings, in which time or space or both are modeled as a discrete or a continuous process. In our application to modeling the multivariate outcome of hospitalization and mortality rates, we target rates across health service areas (HSAs) in the U.S. (regions with relatively self-contained infrastructure for the provision of hospital care) and across time on dialysis. Hence in our application, space is discrete, since HSAs across U.S. are fixed, and time is continuous, since we are interested in drawing inference along the continuous time index of duration on dialysis. While there is extensive literature on multivariate spatiotemporal modeling when space and time are viewed as discrete, the literature on discrete space and continuous time models is limited. When space and time are viewed as discrete, a Markov random field structure in the form of conditionally autoregressive (CAR) or multivariate conditionally autoregressive (MCAR) specifications are commonly used for spatial modeling, while time is typically modeled via an autoregressive structure (see Banerjee et al. (2015) for a comprehensive review). For discrete space and continuous time models, Zhang et al. (2006) proposed a separable structure for the multivariate space-time covariance, while Hepler et al. (2021) and Baer et al. (2021) considered nonseparable space-time covariances. However, both of the latter approaches use quite restricted linear or user-defined parametric forms in temporal modeling.

We propose multivariate spatiotemporal functional principal components analysis (MST-FPCA) to study the joint spatiotemporal patterns of hospitalization and mortality rates among U.S. dialysis patients. The proposed model is based on a multivariate Karhunen-Loéve expansion which models temporal trends nonparametrically via a data-driven lower dimensional multivariate eigenfunction bases. Spatial correlations are induced through a CAR model assumed among region-specific scores, leading to a nonseparable structure on

the multivariate space-time covariance. Functional principal components decompositions for longitudinal and functional processes have been a major modeling theme in the functional data analysis literature, with focus centering on efficient and interpretable representations of functional variability in highly structured settings (Ramsay and Silverman, 2005). More recently, there has been interest in analyzing multiple trajectories, with dependencies among the curves created by spatial or temporal proximity (Crainiceanu et al., 2009; Staicu et al., 2010; Scheffler et al., 2020; Campos et al., 2022). In the multivariate functional setting, several developments rely upon generalizations of the Karhunen-Loève representation (Jacques and Preda, 2014; Chiou et al., 2014), with extensions developed to handle functions observed over different evaluation domains (Happ and Greven, 2018), or to handle multivariate spatiotemporal data in a separable space-time covariance framework (Di Salvo et al., 2015; Ruggieri et al., 2018). However, to date there has been no work proposed to our knowledge on multivariate FPCA that can incorporate spatial correlations in the observed data and model nonseparable multivariate space-time covariance structures.

Similar to developments proposed in Happ and Greven (2018), we propose a computationally efficient estimation procedure for MST-FPCA, which relies only on univariate FPCA expansions. An MCAR structure is induced on the vector of region-specific scores from the univariate FPCA expansions to target the correlation between outcomes in the multivariate response. The estimated between outcome correlation is then incorporated into estimation of the multivariate eigenfunctions. Finally, the region-specific scores and spatial variation parameters are targeted via a Markov chain Monte Carlo (MCMC) framework, conditional on the estimated multivariate eigenfunctions. MST-FPCA achieves computational efficiency by coupling dimension reduction in modeling nonparametric time trends through data-driven lower dimensional multivariate eigenfunctions with modeling of spatial correlations via a parametric CAR structure. Section 1.2 outlines the proposed MST-FPCA model, along with the proposed estimation and inference procedures. Applications of the proposed methodology to data from a large national database, USRDS, allows us to model spatiotemporal trends jointly in hospitalization risk and mortality rates across the U.S., as outlined in Section 1.3. Finite sample properties of MST-FPCA are studied in simulations

of Section 1.4, followed by a brief discussion given in Section 1.5.

1.2 Proposed Multivariate Spatiotemporal Functional Principal Component Analysis (MST-FPCA)

1.2.1 Model Specification

Let $i = 1, \ldots, n$ index regions, $k = 1, \ldots, T$ index time (months) after transition to dialysis and $j = 1, \ldots, J$ index the different dimensions of the J-dimensional outcome vector, $\boldsymbol{X}_i(t_k) = \{X_i^{(1)}(t_k), \ldots, X_i^{(J)}(t_k)\}^\top$. In our application to USRDS data, our multivariate outcome contains region-specific monthly hospitalization and mortality rates (i.e. $J = 2$), where region-specific rates are obtained as averages of dialysis facility-specific rates. Hospitalization and mortality rates at the dialysis facility level are defined as the ratio of the total number of patient hospitalizations or deaths to the total patient follow-up time for that specific facility at month k, respectively. We multiplied the rates by 12 so that the rate unit can be interpreted as a rate per person-year (PPY) consistent with annual national reporting from the USRDS. In our application to USRDS data, region units are taken to be HSAs across U.S. and region-specific rates are analyzed for a total of 24 months (i.e., 2 year follow-up) after transitioning to dialysis. Note that we opt to model the multivariate response as continuous data similar to the works of Li et al. (2021) and amenable to the proposed FPCA framework.

To study the spatiotemporal variation in hospitalization and mortality rates jointly, the proposed MST-FPCA decomposes the J-dimensional multivariate response vector $\boldsymbol{X}_i(t) = \{X_i^{(1)}(t), \ldots, X_i^{(J)}(t)\}^\top$ $(t \in \mathcal{T})$ by

$$\boldsymbol{X}_i(t) = \boldsymbol{\mu}(t) + \sum_{\ell=1}^{\infty} \rho_{i\ell} \boldsymbol{\psi}_\ell(t) + \boldsymbol{\epsilon}_i(t). \tag{1.1}$$

In (1.1), $\boldsymbol{\mu}(t) = \{\mu^{(1)}(t), \ldots, \mu^{(J)}(t)\}^\top$ denotes the overall mean function, $\boldsymbol{\psi}_\ell(t) = \{\psi_\ell^{(1)}(t), \ldots, \psi_\ell^{(J)}(t)\}^\top$ denotes the multivariate eigenfunctions, $\rho_{i\ell}$ are region-specific principal component

(PC) scores, and $\boldsymbol{\epsilon}_i(t) = \{\epsilon_i^{(1)}(t), \ldots, \epsilon_i^{(J)}(t)\}^\top$, $\epsilon_i^{(j)}(t) \sim_{ind} N(0, \sigma_j^2)$, denotes the measurement error. The multivariate eigenfunctions form an orthonormal system, i.e.,

$$\langle\langle \psi_\ell(t), \psi_{\ell'}(t)\rangle\rangle := \sum_{j=1}^{J} \langle \psi_\ell^{(j)}(t), \psi_{\ell'}^{(j)}(t)\rangle_2 = \sum_{j=1}^{J} \int_{\mathcal{T}} \psi_\ell^{(j)}(t)\psi_{\ell'}^{(j)}(t)\mathrm{d}t = \delta_{\ell\ell'},$$

where $\delta_{\ell\ell'} = 1$ for $\ell = \ell'$ and $\delta_{\ell\ell'} = 0$ otherwise. Under the classical multivariate FPCA framework (Ramsay and Silverman, 2005; Happ and Greven, 2018), the multivariate PC scores $\{\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\top : \ell = 1, 2, \ldots\}$, are assumed to be uncorrelated, with zero means and $\mathrm{Var}(\boldsymbol{\rho}_{i\ell}) = \lambda_\ell$, where λ_ℓ denotes the eigenvalues. Similar to univariate FPCA, the multivariate eigenfunctions describe directions of leading modes of variation in the different dimensions of the functional response, while the eigenvalues quantify the amount of variation explained along the identified modes of variation. In practice, the expansion is truncated to include L eigencomponents based on the fraction of variance explained (FVE), where 2 to 3 eigencomponents are retained in most applications, leading to effective dimension reduction of the high-dimensional data.

Different from multivariate FPCA, we induce a CAR structure on the region-specific PC scores $\rho_{i\ell}$ to capture dependencies due to facility- and region-specific practice patterns and infrastructure. Specifically, let the $n \times n$ adjacency matrix $W = \{w_{ii'}\}$ describe the neighborhood structure of regions, where $w_{ii'} = 1$ if regions i and i' $(i \neq i')$ are neighbors, denoted by $i \sim i'$, and $w_{ii'} = 0$ otherwise. By convention, the diagonal elements of W are set to zero. Further let D be the diagonal matrix consisting of elements $d_i = \sum_{i\sim i'} w_{ii'}$, denoting the total number of neighbors of region i. Then the full conditional distribution for the ℓth PC score for region i, $\rho_{i\ell}$, is specified as a weighted average of the ℓth PC scores from neighbors of region i, via the Markov Random Field $\rho_{i\ell}|\{\rho_{i'\ell}\}_{i\neq i'} \sim N(\nu \sum_{i'\sim i} w_{ii'}\rho_{i'\ell}/d_i, \alpha_\ell/d_i)$ with a variance component α_ℓ and a spatial correlation parameter ν. Through Brook's lemma, the joint distribution of the ℓth PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\top$ takes the form $\boldsymbol{\rho}_\ell \sim N(\mathbf{0}, \alpha_\ell(D - \nu W)^{-1})$, where the spatial correlation parameter ν is constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues of the matrix $D^{-1/2}WD^{-1/2}$ in order for the precision matrix $(D - \nu W)/\alpha_\ell$ to be positive definite (Banerjee et al., 2015). The

CAR model induced on the PC scores stabilizes estimation especially for smaller regions, smoothing out scores across neighbors. Note that even though the multiple outcomes considered share the same PC scores in the proposed modeling, this does not imply that the outcomes also share the same spatiotemporal correlation structure. Since the eigenfunctions are allowed to vary across outcomes (through the multivariate eigenfunctions), the proposed modeling is able to induce different spatiotemporal correlations across outcomes (via the linear combination of PC scores multiplying multivariate eigenfunctions) that are nonseparable.

1.2.2 Estimation and Inference

Similar to Happ and Greven (2018), the proposed MST-FPCA is also built on links to univariate FPCA of the different dimensions of the multivariate response vector for computational efficiency. The proposed estimation algorithm begins by expansions of the different dimensions of $\boldsymbol{X}_i(t)$ employing univariate FPCA (see Step 1 in Estimation Algorithm). The dependencies among the vector of univariate PC scores are modeled next, via an MCAR correlation structure, and used in targeting the multivariate eigenfunctions. Specifically, the variation in univariate PC scores is decomposed into a between eigencomponent (Σ_b) and a within eigencomponent (Σ_w) variation, capturing dependencies between the eigencomponents of the different dimensions of the response and spatial correlations across regions within eigencomponents, respectively (Step 2). The multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t)$ are targeted using the estimated univariate eigenfunctions and eigenvectors of the estimated Σ_b (Step 3). Hence, by utilizing the dependency between the eigencomponents of the univariate response (through Σ_b) in deriving the multivariate eigenfunctions, MST-FPCA avoids multivariate FPCA on the higher dimensional covariance processes, which can be quite cumbersome in higher dimensions. Conditional on the estimated mean and multivariate eigenfunctions the multivariate PC scores $\rho_{i\ell}$ and spatial correlation parameters α_ℓ and ν are targeted within a hierarchical modeling framework via MCMC (Step 4). Finally, the region-specific multivariate response trajectories are reconstructed with pointwise credible intervals, relying on the estimated multivariate eigenfunctions and posterior distribution of the multivariate PC

scores obtained in Step 4.

Direct estimation of the multivariate eigenfunctions in the first three steps, instead of expansion on a known bases set has two advantages. First, it achieves dimension reduction, in the sense that in most applications a small number of multivariate eigenfunctions are enough to capture most of the variation in the data, leading to computational savings in the Bayesian modeling of Step 4. In addition and perhaps more importantly, the multivariate eigenfunctions provide extra interpretations in model building. Because they are estimated from the data, they represent dominant modes of variation in time. Hence in the case of the multivariate eigenfunctions, they lead to the study and comparison of modes of temporal variation across outcomes.

Note that there are multiple computational savings utilized in the proposed estimation algorithm that makes the implementation of the proposed methodology feasible for decompositions of high-dimensional multivariate response. First, the proposed algorithm relies only on univariate FPCA decompositions. Second, the MCAR model induced on the univariate PC score vector is modeled as a linear combination of independent latent Gaussian processes with a CAR correlation structure. This lower dimensional representation allows for the efficient fitting of the MCAR model in WinBUGS. The hierarchical model with a CAR correlation structure on the multivariate PC scores in Step 4 is also implemented in WinBUGS, leading to easy implementation of the proposed algorithm. The R code and a tutorial for implementing MST-FPCA are made publicly available on Github (https://github.com/qqian-biostat/MST-FPCA). The proposed estimation algorithm is outlined in the table below, with specific steps discussed in further detail in this section.

Step 1: The proposed MST-FPCA begins by univariate FPCA employed in each dimension $X_i^{(j)}(t)$ of the multivariate outcome. A penalized spline smoother is used to obtain the univariate mean function $\widehat{\mu}^{(j)}(t)$ (where the smoothing parameter can be selected via generalized cross validation). The estimated mean is used to center the observed data, $\widehat{X}_i^{c(j)}(t) = X_i^{(j)}(t) - \widehat{\mu}^{(j)}(t)$, leading to the empirical covariance, $\widehat{G}^{(j)}(t,t') = \sum_{i=1}^{n} \widehat{X}_i^{c(j)}(t)\widehat{X}_i^{c(j)}(t')/n$. Following common practice (Ramsay and Silverman, 2005), the diagonal entries of $\widehat{G}^{(j)}(t,t')$ are removed before the covariance surface is smoothed using two-dimensonal penalized

Estimation Algorithm

Step 1: Employ univariate FPCA in each dimension $(X_i^{(j)}(t))$ of the multivariate response vector $\boldsymbol{X}_i(t)$ to target the estimated univariate eigenfunctions and univariate PC score vector $\boldsymbol{\xi}$.

Step 2: Decompose dependencies among $\widehat{\boldsymbol{\xi}}$ into a between eigencomponent Σ_b and a within eigencomponent Σ_w (between region) variation via an MCAR correlation structure.

Step 3: Target the multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t)$ using the estimated univariate eigenfunctions and eigenvectors of the estimated Σ_b.

Step 4: Fixing the estimated mean and multivariate eigenfunctions, target the multivariate PC scores $\rho_{i\ell}$ and spatial correlation parameters α_ℓ and ν within a hierarchical modeling framework via MCMC.

Step 5: Provide inference on the multivariate response trajectories using the estimated multivariate eigenfunctions and posterior distribution of the multivariate PC scores obtained in Step 4.

smoothing splines. This is because the i.i.d measurement error inflates the error along the diagonal of the covariance. In addition, the smoothing parameters are selected by restricted maximum likelihood (REML) (Li et al., 2021). Once the covariance operators are obtained, estimators of the univariate eigenfunctions $\{\widehat{\phi}_m^{(j)}(t) : m = 1, \ldots, M_j\}$, and PC scores $\{\widehat{\xi}_{im}^{(j)} : i = 1, \ldots, n;\ m = 1, \ldots, M_j\}$ are recovered by FPCA. The numbers of eigencomponents, M_j, retained in the univariate FPCA expansions are determined by the FVE, where we use FVE $> 99\%$ in numerical applications to retain enough information at the initial step of the algorithm.

Step 2: Next, the vector of univariate PC scores $\widehat{\boldsymbol{\xi}}_i^\top = (\widehat{\xi}_{i1}^{(1)}, \ldots, \widehat{\xi}_{iM_1}^{(1)}, \ldots, \widehat{\xi}_{i1}^{(J)}, \ldots, \widehat{\xi}_{iM_J}^{(J)})$ are modeled via an MCAR model. Let $\Xi = [\widehat{\boldsymbol{\xi}}_1, \ldots, \widehat{\boldsymbol{\xi}}_n]^\top$ denote the $n \times M^+$ score matrix with $M^+ = M_1 + M_2 + \ldots + M_J$ denoting the total number of eigencomponents retained across all J dimensions. Stacking the columns of Ξ leads to the $nM^+ \times 1$ score vector $\widehat{\boldsymbol{\xi}} = \text{vec}(\Xi)$ that is modeled as a linear combination of M^+ $n \times 1$ independent latent spatial Gaussian processes $\mathbf{f}_\ell \sim N(\mathbf{0}, (D - \nu W)^{-1})$, $\ell = 1, \ldots, M^+$, as in Jin et al. (2007), where D and W are as defined in Section 1.2.1 and ν denotes the common spatial smoothing parameter,

$$\underset{nM^+\times 1}{\widehat{\boldsymbol{\xi}}} = \underset{nM^+\times 1}{\boldsymbol{\xi}} + \underset{nM^+\times 1}{\mathbf{e}} = (\underset{M^+\times M^+}{A} \otimes \underset{n\times n}{I}) \underset{nM^+\times 1}{\mathbf{f}} + \underset{nM^+\times 1}{\mathbf{e}}. \tag{1.2}$$

In (1.2), $\mathbf{f} = (\mathbf{f}_1^\top, \ldots, \mathbf{f}_{M^+}^\top)^\top$ denotes the vector of stacked latent Gaussian processes with

the joint distribution $\mathbf{f} \sim N(\mathbf{0}, I_{M^+} \otimes (D - \nu W)^{-1})$, $A = \{a_{\ell\ell'}\}$, $1 \leq \ell' \leq \ell \leq M^+$, denotes an $M^+ \times M^+$ full rank lower triangular matrix and $\mathbf{e} \sim N(\mathbf{0}, \tau^2 I_{nM^+})$ denotes the vector of measurement errors, assumed to be uncorrelated with $\boldsymbol{\xi}$. With the proposed specification in (1.2), the covariance of $\boldsymbol{\xi}$, denoted by Σ, can be decomposed into a between eigencomponent (Σ_b) and a within eigencomponent (Σ_w) variation,

$$\begin{aligned} \underset{nM^+ \times nM^+}{\Sigma} &= (\underset{M^+ \times M^+}{A} \otimes \underset{n \times n}{I}) \text{Bdiag}\{\underset{n \times n}{(D - \nu W)^{-1}}, \ldots, (D - \nu W)^{-1}\}(A^\top \otimes I) \\ &= \underset{M^+ \times M^+}{AA^\top} \otimes \underset{n \times n}{(D - \nu W)^{-1}} \equiv \underset{M^+ \times M^+}{\Sigma_b} \otimes \underset{n \times n}{\Sigma_w}, \end{aligned}$$

where $\Sigma_b = AA^\top$, $\Sigma_w = (D - \nu W)^{-1}$ and Bdiag$(\cdot)$ denotes a block-diagonal matrix. Hence, an MCAR (ν, Σ_b) structure is induced on $\boldsymbol{\xi}$. Representation of $\boldsymbol{\xi}$ as a linear combination of independent spatial Gaussian processes modeled with a CAR structure, via the lower triangular matrix A, allows for easy implementation of the MCAR model in WinBUGS, without the computational burden of operations with large covariance matrices and without the need to check for positive-definiteness of covariance matrices in each iteration of the algorithm. While the between eigencomponent variation captures the dependency between the eigencomponents from univariate expansions, the within eigencomponent (between region) variation captures the spatial dependency among the regional units. Note that while a separable correlation structure is induced on the univariate PC scores, this does not imply a separable correlation structure between time and space, nor the same level spatial smoothing for the multivariate outcome vector. The goal of the modeling of $\boldsymbol{\xi}$ through an MCAR structure with a common between eigencomponent dependence across regions is to incorporate the estimated between eigencomponent dependency into estimation of the multivariate eigenfunctions in Step 3.

The parameters of the MCAR model are targeted via MCMC. Elementwise priors are imposed on the lower-triangular matrix A: $a_{\ell\ell} \sim \text{Lognormal}(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}})$ and $a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$ for $1 \leq \ell' < \ell \leq M^+$. In addition, an Inverse Gamma (IG) (a_{τ^2}, b_{τ^2}) prior is imposed on the measurement error variance τ^2, and a Uniform prior is used for the spatial parameter ν with the parameters constrained to lie between bounds given by the inverse of the minimum and

maximum eigenvalues of the matrix $D^{-1/2}WD^{-1/2}$ (denoted by a_ν, b_ν, respectively). The posterior distribution we seek can be expressed as

$$\pi(a_{\ell\ell}, a_{\ell\ell'}, \nu, \tau^2|\widehat{\boldsymbol{\xi}}) \propto N(\widehat{\boldsymbol{\xi}}|\boldsymbol{\xi}, \tau^2) \times N\left(\boldsymbol{\xi}|\mathbf{0}, AA^\top \otimes (D - \nu W)^{-1}\right) \times \prod_{\ell=1}^{M^+} \text{Lognormal}(a_{\ell\ell}|\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}}),$$
$$\times \prod_{\ell=1}^{M^+} \prod_{\ell'=1}^{\ell-1} N(a_{\ell\ell'}|\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}}) \times \text{Unif}(\nu|a_\nu, b_\nu) \times \text{IG}(\tau^2|a_{\tau^2}, b_{\tau^2}).$$

The model parameters are sampled from the posterior distributions using MCMC with Gibbs sampling and random walk metropolis as implemented in WinBUGS.

Step 3: The between eigencomponent dependence targeted via Σ_b is incorporated in estimation of the multivariate eigenfunctions as

$$\widehat{\psi}_\ell^{(j)}(t) = \sum_{m=1}^{M_j} [\widehat{\mathbf{c}}_\ell]_m^{(j)} \widehat{\phi}_m^{(j)}(t), \quad \ell = 1, \ldots, M^+, \quad j = 1, \ldots, J, \tag{1.3}$$

where each dimension of the multivariate eigenfunction is targeted as a linear combination of the univariate eigenfunctions estimated for that dimension with weights $[\widehat{\mathbf{c}}_\ell]_m^{(j)}$ equal to the mth entry of $[\widehat{\mathbf{c}}_\ell]^{(j)} \in \mathbb{R}^{M_j}$, the jth block of the ℓth eigenvector $\widehat{\mathbf{c}}_\ell$ of $\widehat{\Sigma}_b$ (the posterior mean of Σ_b). The form in (1.3) follows from the fact that the variance components α_ℓ are also the eigenvalues of the between eigencomponent covariance matrix Σ_b of the univariate PC scores as shown in Appendix A.1. We defer the readers to Appendix A.1 for more details.

Step 4: Once the multivariate mean and eigenfunctions are estimated, we target the region-specific PC scores ($\rho_{i\ell}$), spatial parameters (α_ℓ and ν) and measurement error variance (σ_j^2) of MST-FPCA, conditional on $\widehat{\boldsymbol{\theta}} = \{\widehat{\boldsymbol{\mu}}(t), \widehat{\boldsymbol{\psi}}_\ell(t), M^+\}$, via the multivariate hierarchical model

$$\boldsymbol{X}_i(t) = \widehat{\boldsymbol{\mu}}(t) + \sum_{\ell=1}^{M^+} \rho_{i\ell} \widehat{\boldsymbol{\psi}}_\ell(t) + \boldsymbol{\epsilon}_i(t),$$

$$\underset{n\times 1}{\boldsymbol{\rho}_\ell} = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\top \sim N(\underset{n\times 1}{\mathbf{0}}, \alpha_\ell(\underset{n\times n}{D} - \nu \underset{n\times n}{W})^{-1}), \quad \underset{J\times 1}{\boldsymbol{\epsilon}_i(t)} \sim N(\underset{J\times 1}{\mathbf{0}}, \underset{J\times J}{\text{Diag}(\sigma_1^2, \ldots, \sigma_J^2)}),$$

$$\sigma_j^2 \sim \text{IG}(a_{\sigma_j^2}, b_{\sigma_j^2}), \quad \alpha_\ell \sim \text{IG}(a_{\alpha_\ell}, b_{\alpha_\ell}), \quad \nu \sim \text{Unif}(a_\nu, b_\nu),$$

where $\text{Diag}(\sigma_1^2, \ldots, \sigma_J^2)$ denotes the $J \times J$ diagonal matrix with σ_j^2, $j = 1, \ldots, J$, on the diagonal (Step 4). Hence the posterior distribution we seek can be expressed as

$$\begin{aligned}\pi(\boldsymbol{\rho}_\ell, \nu, \alpha_\ell, \sigma_j^2 | \boldsymbol{X}_i(t), \widehat{\boldsymbol{\theta}}) \propto & \prod_{i=1}^{n} \prod_{k=1}^{T} N\left(\boldsymbol{X}_i(t_k) | \{\widehat{\boldsymbol{\mu}}(t_k) + \sum_{\ell=1}^{M^+} \rho_{i\ell} \widehat{\boldsymbol{\psi}}_\ell(t_k)\}, \text{Diag}(\sigma_1^2, \ldots, \sigma_J^2) \right) \\ & \times \prod_{\ell=1}^{M^+} N(\boldsymbol{\rho}_\ell | \mathbf{0}, \alpha_\ell (\text{D} - \nu \text{W})^{-1}) \times \prod_{\ell=1}^{M^+} \text{IG}(\alpha_\ell | a_{\alpha_\ell}, b_{\alpha_\ell}) \\ & \times \text{Unif}(\nu | a_\nu, b_\nu) \times \prod_{j=1}^{J} \text{IG}(\sigma_j^2 | a_{\sigma_j^2}, b_{\sigma_j^2}).\end{aligned}$$

The model parameters are sampled from the posterior distributions using MCMC as implemented in WinBUGS. Since M^+ is the total number of univariate eigencomponents retained across the J dimensions, the total number (L) of multivariate eigencomponents used in MST-FPCA could be chosen to be smaller than M^+ in applications where M^+ may be large. Here we recommend the use of the estimated variance components α_ℓ in formulating the fraction of variance explained (FVE), since they are established as eigenvalues of the between eigencomponent covariance matrix Σ_b (see Appendix A.1 for details).

Step 5: In the final step, region-specific trajectories are constructed for the multivariate outcome using the estimated MST-FPCA model components: $\widehat{\boldsymbol{X}}_i(t) = \widehat{\boldsymbol{\mu}}(t) + \sum_{\ell=1}^{L} \widehat{\rho}_{i\ell} \widehat{\boldsymbol{\psi}}_\ell(t)$. In addition, $(1-\gamma)$ pointwise credible intervals (CIs) for region-specific trajectories are obtained by

$$\widehat{\boldsymbol{X}}_i(t) \pm \Phi^{-1}(1 - \frac{\gamma}{2}) \sqrt{\text{diag}\left[\text{Var}\{\widehat{\boldsymbol{X}}_i(t) - \boldsymbol{X}_i(t) | \widehat{\boldsymbol{\theta}}\} \right]},$$

where $\text{Var}\{\widehat{\boldsymbol{X}}_i(t) - \boldsymbol{X}_i(t) | \widehat{\boldsymbol{\theta}}\} \approx \widehat{\Psi} \widehat{\Omega} \widehat{\Psi}^\top$, with $\widehat{\Psi} = \{\widehat{\boldsymbol{\psi}}_1(t), \ldots, \widehat{\boldsymbol{\psi}}_L(t)\}$, $\Phi(\cdot)$ denotes the Gaussian cumulative distribution function, $\text{diag}(\cdot)$ denotes the diagonal of a matrix and $\Omega_{L \times L} = \text{Var}\{\boldsymbol{\rho}_i | \boldsymbol{X}(t), \widehat{\boldsymbol{\theta}}\}$ for $\boldsymbol{\rho}_i = (\rho_{i1}, \ldots, \rho_{iL})^\top$ is targeted using the posterior variance of the estimated multivariate PC scores.

When the different dimensions of the multivariate outcome have different domains/ranges or if they exhibit different amounts of variation, standardization of the data as a preliminary step may be necessary in order to obtain interpretable multivariate functional principal components (Jacques and Preda, 2014; Chiou et al., 2014; Happ and Greven, 2018).

The standardization can be carried out by rescaling data in each dimension using weights $s_j = \left\{\int_{\mathcal{T}} \widehat{\text{Var}}\left(X^{(j)}(t)\right) \mathrm{d}t\right\}^{-1}$, such that the integrated variance along the rescaled data $\tilde{X}^{(j)}(t) = s_j^{1/2} X^{(j)}(t)$ equals one (Happ and Greven, 2018). In this way all dimensions of the multivariate outcome contribute equal amounts of variation to the analysis, similar to the standardization of classical multivariate PCA. In our data application, this rescaling is utilized. Note that when standardization is employed as a preliminary step in MST-FPCA, the algorithm uses the rescaled data in the first three steps in deriving the multivariate eigenfunctions. Once the multivariate eigenfunctions are estimated, the multivariate hierarchical model of Step 4 and inference in Step 5 utilizes the observed data as outlined in this section.

1.3 Multivariate Spatiotemporal Functional Principal Component Analysis in the Dialysis Population

The USRDS collects data on nearly all patients with ESKD in the U.S. Our study cohort includes patients aged 18 years or older who transitioned to dialysis between January 1, 2005, and September 30, 2013. Patients were followed for a maximum of two years, beginning on day 91 of dialysis (after 90 days to establish a stable treatment modality), with the last date of follow-up on December 31, 2015. Facility-specific hospitalization and mortality rates per person-year (PPY) are calculated monthly over the two year follow-up and are averaged within regions to yield the region-specific multivariate outcome. Consistent with national USRDS reporting, regions are taken to be Health Service Areas (HSAs), which are regions with relatively self-contained infrastructure for the provision of hospital care in the contiguous U.S., including the District of Columbia. The final study cohort contains 367 regions/HSAs, where we defer detailed descriptions of the study cohort, exclusion rules, and preprocessing steps to Appendix A.2. The mean region-specific hospitalization and mortality rates are 1.810 and 0.073 PPY, respectively, where the raw hospitalization and mortality rate trajectories over the two-year follow-up are given in Figure 1.1 (a, c). Note that the range of the two outcomes are quite different and hence a preliminary rescaling is applied with $s_1 = 0.40$ and $s_2 = 215.07$, for hospitalization and mortality, respectively, as discussed in

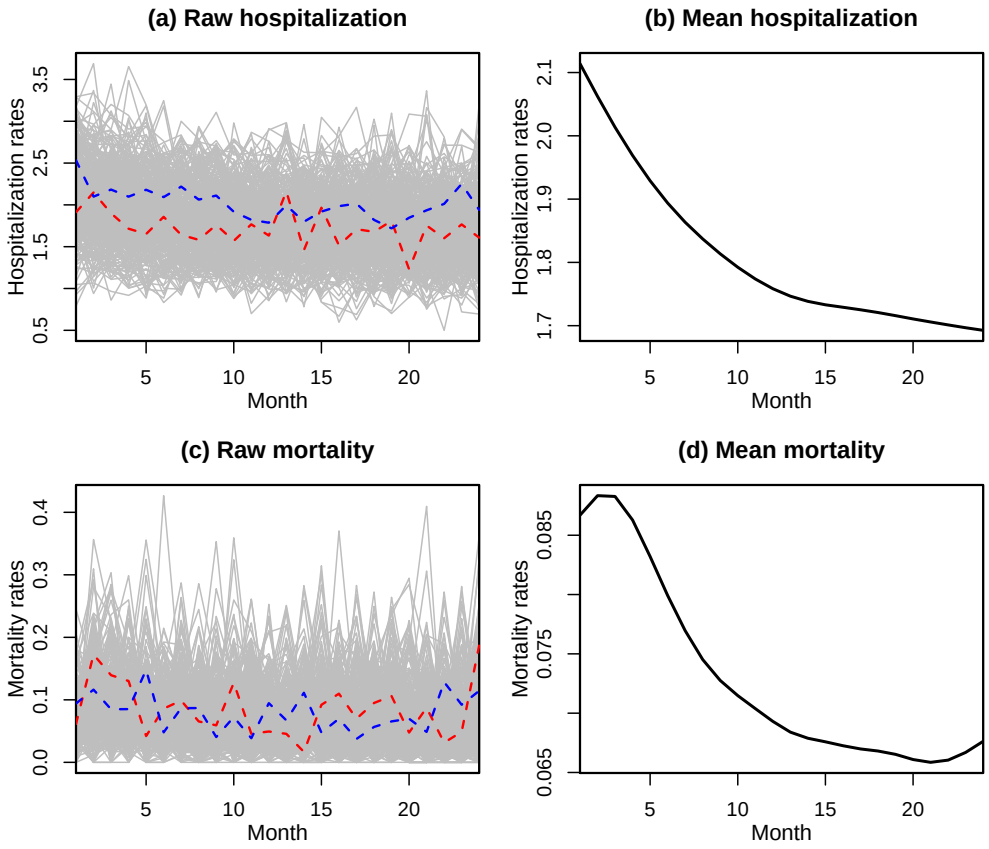

Figure 1.1: Raw region-specific hospitalization (a) and mortality (c) trajectories. Raw trajectories for a randomly chosen two regions are depicted with dashed and dotted lines. The estimated mean trajectories for hospitalization and mortality are provided in (b) and (d), respectively.

Section 1.2.2 before MST-FPCA, to guarantee that variation along both dimensions of the multivariate outcome is captured in the estimated multivariate eigenfunctions.

1.3.1 Estimated MST-FPCA Components

The mean hospitalization and mortality trajectories are given in Figures 1.1 (b, d), where the highest rates in both outcomes are observed within three months after transitioning to dialysis and where both (average) rates steadily decline during the first year on dialysis, consistent with previous literature (Foley et al., 2014). For example, the mean hospitalization rate at 1 month is 2.114 PPY and declines 16.84% to 1.758 PPY by 12 months; and similarly, the mean mortality rate at 1 month is 0.087 PYY and decreases to 0.069 PPY by 12 months

(20.69% decline). While hospitalization rates continue to decline after the first year, the mortality rates remain relatively stable in the second year on dialysis. A total of four and two eigencomponents are retained in univariate FPCA expansions of hospitalization and mortality, respectively, explaining 99.2% and 99.3% of the total variation. The estimated spatial correlation parameter $\widehat{\eta}$ equals 0.923, leading to the within eigencomponent correlations in the range of 0.27 to 0.63 in neighboring HSAs. The estimated multivariate eigenfunctions with respective estimated spatial variance parameters $\widehat{\alpha}_1 = 3.090$, $\widehat{\alpha}_2 = 1.966$, $\widehat{\alpha}_3 = 0.211$, $\widehat{\alpha}_4 = 0.117$, $\widehat{\alpha}_5 = 0.103$, and $\widehat{\alpha}_6 = 0.065$, are given in Figure A.1 in Appendix A. The two leading multivariate eigenfunctions (given in Figure A.1 (a)-(b)) describe mostly constant variation in both hospitalization and mortality rates, with slightly higher variation in mortality in the first year of dialysis. The third, fourth and fifth multivariate eigenfunctions (given in Figure A.1 (c)-(d)) mainly explain variation in hospitalization, which highlight variation in the first and last six months on dialysis (third eigenfunction), followed by variation in the middle one year and end of the two year follow-up (fourth eigenfunction) and finally at one year and 18 months on dialysis (fifth eigenfunction). The last multivariate eigenfunction (given in Figure A.1 (e)-(f)) mainly explains variation in mortality, which highlights variation in mortality at initiation and one year and 18 months after transitioning to dialysis. The leading time-varying variation in hospitalization observed within the first six months of dialysis is consistent with higher hospitalization rates observed at initiation of dialysis, while higher variation in the last six months of follow-up may be related to the decrease in the total number of patients towards the end of the two year follow-up.

1.3.2 Inference for Region-Specific Hospitalization and Mortality Rates

The six estimated multivariate eigenfunctions are retained in the multivariate reconstruction of the region-specific hospitalization and mortality trajectories. The raw and predicted hospitalization and mortality rates from the 3rd, 12th, and 18th months on dialysis are displayed for all HSAs in Figure 1.2. There is a distinct pattern ("band") of higher rates (darker blue) in both hospitalization and mortality from Massachusetts to southern Texas. While regions display higher rates of hospitalization more consistently in this band, mortality

rates are more variable with regions of elevated and lower mortality rates. In addition, there are outcome-specific patterns that emerge where some HSAs in northern California, Oregon, Montana, and Idaho have relatively low hospitalization rates but elevated mortality rates, whereas some HSAs in Florida and Arizona have relatively high hospitalization rates but not very high mortality rates. Furthermore, consistent with the estimated mean functions for hospitalization and mortality displayed in Figures 1.1 (b) and 1.1 (d), respectively, the hospitalization and mortality rates are the highest in the first few months on dialysis, where hospitalization rates consistently decrease over the two year follow-up and mortality rates decrease within the first year on dialysis and remain roughly stable in the second year.

For a more in-depth study of the variation in hospitalization and mortality rates across U.S., we divide the regions into five major zones: West (41 HSAs), Midwest (98 HSAs), Southwest (41 HSAs), Southeast (141 HSAs), and Northeast (46 HSAs). Table A.1 in Appendix A displays the median and the 5th and 95th percentiles of the predicted region-specific hospitalization and mortality rates at selected months 3, 12, and 18 across the five zones and Figure 1.3 displays the median rates. In addition, consistent with the estimated mean functions for hospitalization and mortality in Figures 1.1 (b) and 1.1 (d), respectively, both overall hospitalization and mortality rates decrease at a faster rate in the first year on dialysis, where they continue to decrease in the second year as well across the five zones but at a slower rate than the first year. More specifically, the median hospitalization rates decline from 3 to 18 months on dialysis by 0.184 to 0.362 PPY across five zones, which represents a 11.65%, 13.32%, 14.25%, 14.89% and 16.41% decline in the West, Midwest, Southwest, Southeast, and Northeast, respectively. Similarly, the mortality rates decline in the first year on dialysis from 3 to 12 months by 0.010 to 0.022 PPY across five zones, which represents a 14.93%, 25.00%, 24.39%, 22.47% and 18.82% decline in the West, Midwest, Southwest, Southeast, and Northeast, respectively (See Table A.1).

Figure A.2 depicts the predicted hospitalization and mortality trajectories for five HSAs with median rates from the five zones. The predicted trajectories for HSAs from the five zones are plotted (in red) along with their 95% pointwise credible intervals (CI), overlaying predictions for a representative HSA with median rates across the entire U.S. (in blue) for

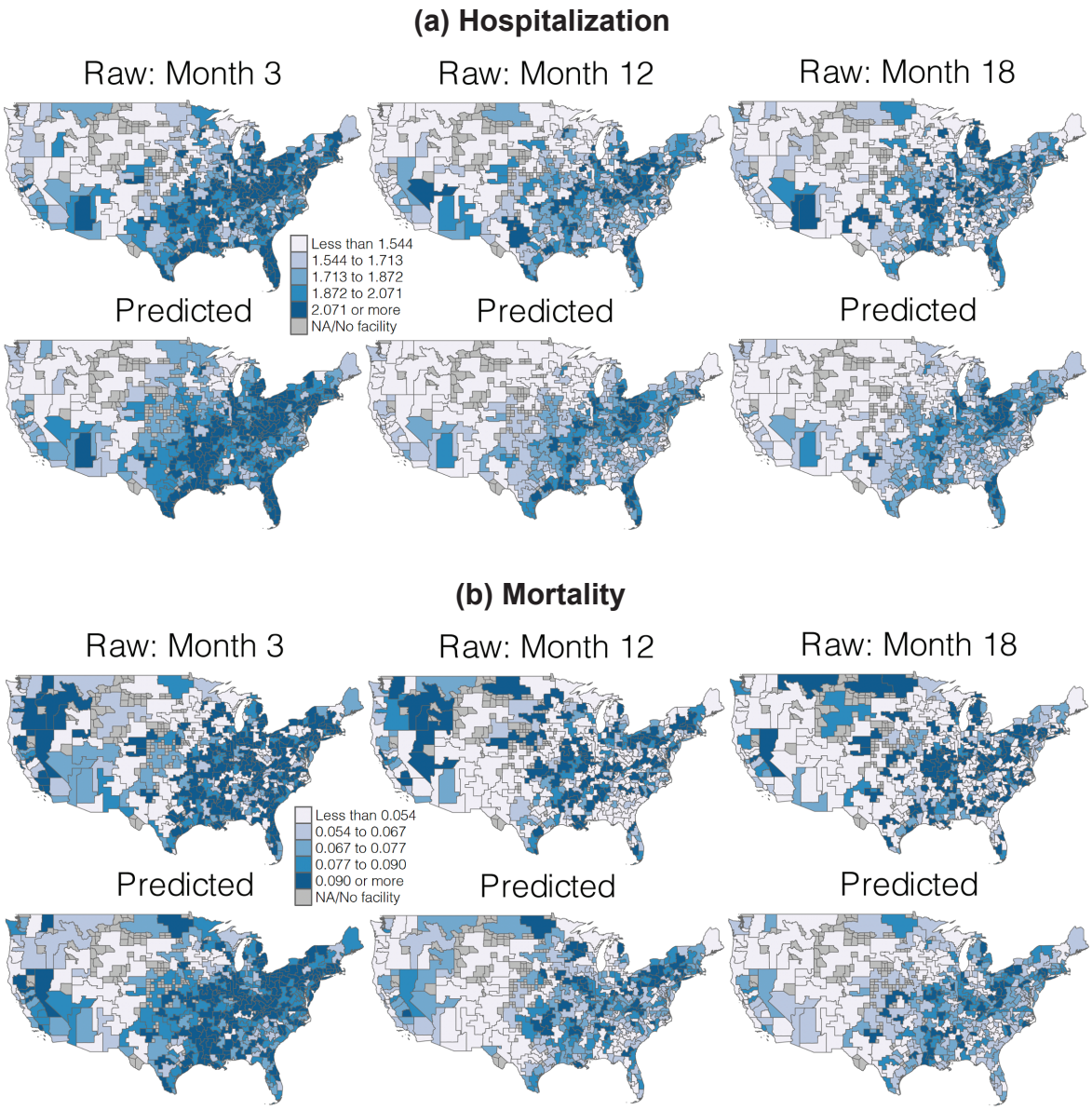

Figure 1.2: The raw and predicted hospitalization (a) and mortality (b) rates from the 3rd, 12th, and 18th months on dialysis for all HSAs.

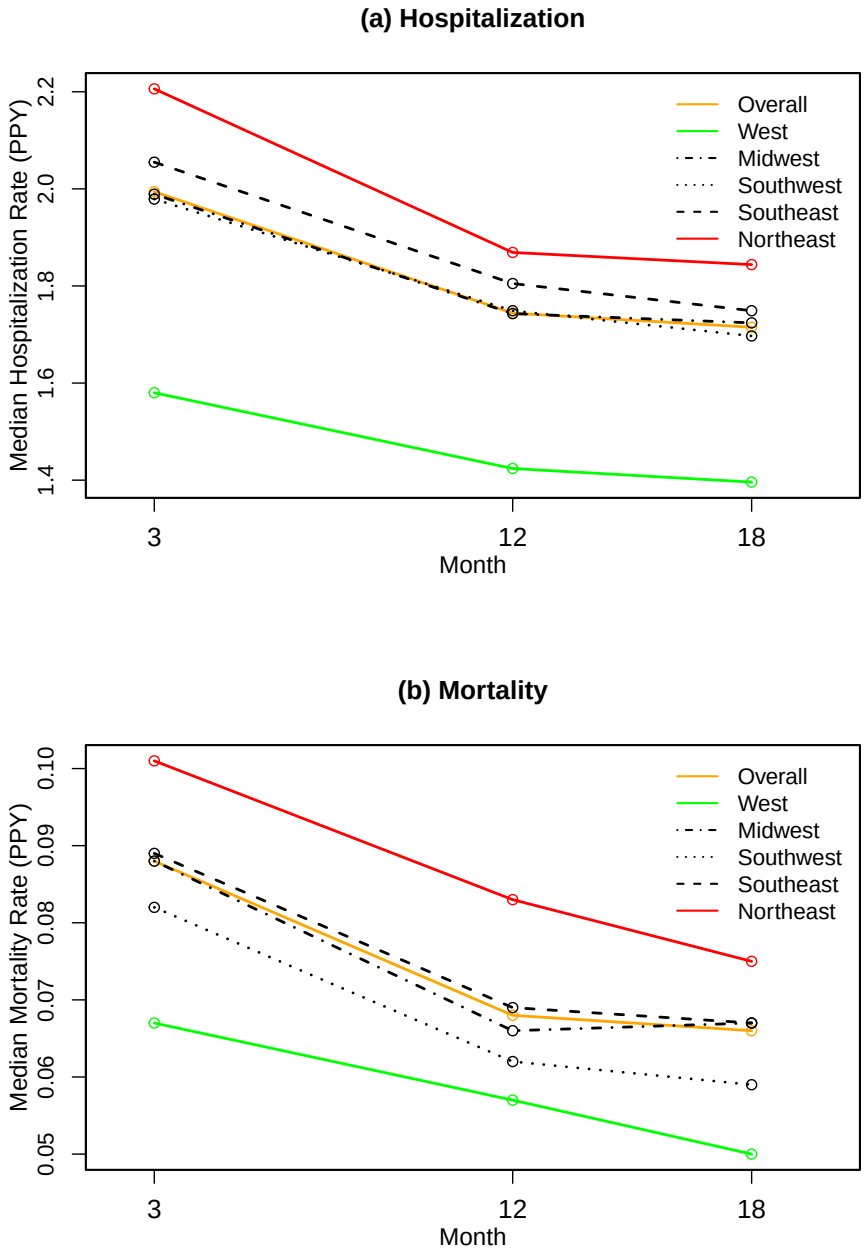

Figure 1.3: Median predicted hospitalization (a) and mortality (b) rates per person-year (PPY) for five major zones, as well as the overall rates across all health service areas (HSAs).

comparison. Consistent with previous observations, hospitalization rates are the lowest in the West, where the CI for rates from the representative HSA does not overlap with the CI for the HSA from the West for most of the two year follow-up. While mortality rates are still the lowest in the West, the CIs overlap indicating more variability in mortality and lack of significant separation in the rates. This also aligns with the previous observation that some HSAs in the West (northern California) have higher mortality but lower hospitalization rates (Figure 1.2). In addition, the Northeast has the highest hospitalization and mortality rates, consistently over the two year follow-up where the CIs that compare rates to a representative HSA do not separate. Rates from HSAs from the remaining three zones in the Midwest, Southwest and Southeast are more similar to those from the representative HSA.

To further assess model fit, we study the residuals obtained from MST-FPCA fits and compare the MST-FPCA fits to fits from three alternate methods of modeling the data. The map of absolute value of residuals from the MST-FPCA fits at the 3rd, 12th, and 18th months on dialysis for all HSAs is given in Figure A.3. The residuals are quite small for hospitalizations, with larger deviations observed in less than 5% of HSA in the harder to model outcome of mortality. Most of these larger deviations are observed on the band of higher rates from Massachusetts to southern Texas and correspond to higher than expected mortality (positive residual). In addition, there does not seem to be an obvious spatial correlation in the residuals, implying that MST-FPCA is able to model spatial correlations in the data effectively. The first alternate method MST-FPCA is compared to (referred to as multivariate FPCA (M-FPCA)) assumes that the multivariate PC scores are i.i.d. across regions, ignoring spatial correlations in the data. M-FPCA is fitted through the algorithm of Happ and Greven (2018). The second comparative approach considered is called the Multivariate 'Besag-York-Mollié' model (MBYM) proposed by Boulieri et al. (2017). MBYM utilizes a random walk of order one to induce temporal correlations, and a multivariate intrinsic conditional autoregressive (ICAR) structure to model spatial pattens, leading to a separable discrete-time, discrete-space modeling framework. MBYM is easy to fit via the built-in `mv.car` package in WinBUGS, hence is utilized as a comparative method. The last alternate method considered is to fit univariate spatiotemporal FPCA

separately in each dimension, which we refer to as U-FPCA. Similar to MST-FPCA, the U-FPCA approach targets the univariate eigenfunctions in the first step via FPCA. Then in each dimension, a CAR structure is induced among PC scores to capture dependencies among regions. The detailed description of the three alternate models is deferred to Appendix A.3. We compare the four fits using relative mean squared deviation error (MSDE), i.e., $\text{MSDE}_{\widehat{x}^{(j)}(t)} = \int\{X_i^{(j)}(t) - \widehat{X_i}^{(j)}(t)\}^2 / \int\{X_i^{(j)}(t)\}^2$. The mean MSDEs across regions are (0.0179, 0.0189, 0.0178, 0.0167) and (0.2587, 0.2564, 0.2552, 0.2454) in hospitalization and mortality for (M-FPCA, MBYM, U-FPCA, MST-FPCA), respectively. The run times for the four fits are (6.33, 81.31, 8.77, 12.02) minutes, for (M-FPCA, MBYM, U-FPCA, MST-FPCA), respectively. MST-FPCA and U-FPCA are fit with two parallel chains with 15000 iterations (5000 burn-in) and 3500 iterations (1000 burn-in) in the MCMC involved in Steps 2 and 4, respectively. M-FPCA and MYBM are also fit using two parallel chains with 5000 iterations (1000 burn-in) and 30000 iterations (10000 burn-in), respectively. The MST-FPCA leads to smaller MSDEs in modeling hospitalizations and modest improvements in modeling mortality, where the computational efficiency is pretty close to M-FPCA which ignores the spatial correlations and U-FPCA in lower dimensions. More information on comparisons of MST-FPCA to the three alternate methods in simulation studies are summarized in the next section.

1.4 Simulation Studies

Simulation studies are conducted to examine the finite sample properties of the proposed MST-FPCA, as well as to compare its performance to the performance of M-FPCA, MBYM and U-FPCA. Two simulation set-ups are considered: 1) independent multivariate response and 2) correlated multivariate response, with varying number of time points ($T = 24, 50$), number of regions ($n = 49, 367$), different levels of error variance ($\sigma^2 = 0.02, 0.5$) and multivariate response from two and three dimensions. Simulation results are briefly reviewed here, where we defer details to A.4 and A.5, and more simulation results can be found in Table A.2 – A.5, as well as Figure A.4 – A.6. Error measures for all model components get

smaller with decreasing noise level σ^2. Similarly, all error measures decrease with increasing number of time points and regions, as expected, where the trends are stronger in estimation of multivariate eigenfunctions. Even with smaller number of time points, number of regions, and higher level of error variance, error measures from MST-FPCA signal a good fit (see Table 1.1 and Table A.2).

Table 1.1: The mean MSE and MSDE values for estimation of the model components from the correlated response simulation set-up with varying number of time points T, varying number of regions n and different levels of measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

$\boldsymbol{T}$	**24 time points**				**50 time points**			
$\boldsymbol{n}$	**49 regions**		**367 regions**		**49 regions**		**367 regions**	
σ^2	0.02	0.5	0.02	0.5	0.02	0.5	0.02	0.5
				MSDE				
$\widehat{\psi}_1^{(1)}(t)$	0.079	0.116	0.030	0.033	0.065	0.104	<0.001	0.001
$\widehat{\psi}_1^{(2)}(t)$	0.079	0.115	0.030	0.034	0.065	0.104	<0.001	0.001
$\widehat{\psi}_2^{(1)}(t)$	0.090	0.125	0.035	0.039	0.074	0.105	0.001	0.002
$\widehat{\psi}_2^{(2)}(t)$	0.090	0.127	0.035	0.039	0.074	0.104	0.001	0.002
$\widehat{\psi}_3^{(1)}(t)$	0.010	0.048	<0.001	0.004	0.010	0.030	<0.001	0.002
$\widehat{\psi}_3^{(2)}(t)$	0.010	0.051	<0.001	0.005	0.010	0.030	<0.001	0.003
				MSE				
$\widehat{\alpha}_1$	0.057	0.127	0.047	0.051	0.006	0.030	<0.001	0.002
$\widehat{\alpha}_2$	0.014	0.052	0.007	0.007	0.003	0.045	<0.001	0.002
$\widehat{\alpha}_3$	0.023	0.105	0.021	0.024	0.007	0.085	0.004	0.005
$\widehat{\nu}$	0.146	0.157	0.054	0.059	0.144	0.152	0.019	0.049
$\widehat{\sigma}^2$	< 0.001	0.002	< 0.001	< 0.001	<0.001	<0.001	<0.001	<0.001
$\widehat{\rho}_{i1}$	0.221	0.484	0.050	0.080	0.208	0.475	0.011	0.027
$\widehat{\rho}_{i2}$	0.364	0.392	0.120	0.170	0.316	0.342	0.002	0.027
$\widehat{\rho}_{i3}$	0.080	0.292	0.009	0.140	0.077	0.215	0.009	0.072

The prediction and inference for the multivariate region-specific trajectories based on M-FPCA, MBYM, U-FPCA and MST-FPCA are evaluated using MSDE and the coverage probability (CP) and length of the associated 95% pointwise CIs. Ignoring the spatial correlation in M-FPCA leads to higher MSDEs in recovery of the multivariate trajectories and lower coverage probabilities (compared to MST-FPCA and U-FPCA). In addition, the joint modeling of correlated outcomes via MST-FPCA leads to better efficiency than univari-

Table 1.2: The mean MSDE of the predicted multivariate region-specific trajectories and the coverage probability (CP) and length of the associated 95% pointwise credible intervals (CIs) based on M-FPCA, MBYM, U-FPCA and MST-FPCA from both independent and correlated response simulation set-ups for number of time points $T = 24$, number of regions $n = 367$ and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	**367 regions**							
Noise level, σ^2	0.02				0.5			
Model	M-FPCA	MBYM	U-FPCA	MST-FPCA	M-FPCA	MBYM	U-FPCA	MST-FPCA
	Independent response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.071	1.031	0.016	0.015	0.505	1.048	0.363	0.372
MSDE: $\widehat{x}_i^{(2)}(t)$	0.068	0.515	0.016	0.016	0.478	0.604	0.330	0.336
Length$^{(1)}$	0.194	0.409	0.194	0.194	0.867	0.516	0.870	0.874
Length$^{(2)}$	0.192	0.427	0.192	0.192	0.857	0.664	0.868	0.865
CP$^{(1)}$ (%)	90.11	21.67	94.09	94.08	91.87	27.06	94.53	94.42
CP$^{(2)}$ (%)	90.44	42.62	94.30	94.29	91.11	54.99	94.57	94.05
	Correlated response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.073	1.032	0.016	0.011	0.474	1.048	0.284	0.188
MSDE: $\widehat{x}_i^{(2)}(t)$	0.072	1.031	0.016	0.012	0.471	1.048	0.291	0.190
Length$^{(1)}$	0.157	0.408	0.194	0.179	0.750	0.515	0.877	0.755
Length$^{(2)}$	0.157	0.408	0.194	0.179	0.749	0.514	0.874	0.754
CP$^{(1)}$ (%)	91.13	21.61	94.09	95.13	91.52	27.01	94.56	96.19
CP$^{(2)}$ (%)	90.49	22.64	93.72	94.63	90.68	28.08	94.48	96.06
	Run times							
	M-FPCA		MBYM		U-FPCA		MST-FPCA	
Time(min)	6.330		81.307		8.774		12.022	

ate modeling via U-FPCA, as expected, and the separable covariance structure of MBYM across time and space, coupled with constant spatial correlation assumed across outcomes via ICAR, is too restrictive and leads to worse performance compared to the other comparative methods (see Table 1.2, Table A.3 – A.5).

1.5 Discussion

We proposed a novel multivariate spatiotemporal functional principal component analysis model to study the joint spatiotemporal patterns of hospitalization and mortality rates among dialysis patients in the U.S. The proposed MST-FPCA not only allows for leveraging information from nearby rates (over geographic regions and time), but also across correlated outcomes to achieve a more comprehensive assessment of variation in the multivariate out-

come. The proposed estimation involves effective dimension reduction in modeling variation across time through multivariate eigenfunctions, which provide additional interpretations on directions of dominant temporal variation. Spatial correlations on univariate PC scores are modeled via MCAR. Both the multivariate eigenfunctions and the proposed MCAR modeling are built on computationally feasible representations involving lower dimensional building blocks (univariate eigenfunctions and independent latent Gaussian processes with a CAR correlation structure, respectively). Hence, the proposed modeling, while leading to a nonseparable space and time covariance structure in the outcomes, can still easily scale up to multivariate response in higher dimensions. In addition, the proposed modeling via MCMC allows for estimation and inference on multivariate hospitalization and mortality trajectories in order to obtain regional hot spots (with high rates in both hospitalization and mortality or with differing patterns in the outcomes) and time periods with elevated rates after transitioning to dialysis. Results point to significant spatiotemporal variation in the multivariate outcome across the U.S.

Since the estimated multivariate mean and eigenfunctions are considered fixed in the Bayesian modeling (Step 4), the proposed inference can underestimate uncertainty in small data applications where estimation error in targeting the mean and eigenfunctions are non-negligible. For small data applications, corrected inference has been proposed in functional data settings via parametric bootstrap (Li et al., 2021). A similar extension could enable MST-FPCA to incorporate the stochasticity in estimation of the mean and multivariate eigenfunctions and is the topic for future research. We follow previous works (Quick et al., 2013; Li et al., 2021) in modeling rates directly, amenable to functional data analysis techniques used, however, another extension of MST-FPCA that is of interest is in modeling generalized outcome. Multivariate FPCA has been extended for generalized outcome utilizing a semiparametric latent process (Jiang et al., 2022). Multivariate eigenfunctions can be incorporated into a generalized Bayesian hierarchical framework in such an extension, with a need for a new set of tools for computational savings. Finally, the proposed methodology can be extended for modeling time-varying risk factors that may explain parts of the variation observed. The extension would expand the mean function $\mu(t)$ (and the possibly

time-varying covariate effect functions) on a common basis system whose coefficients can be targeted within the proposed Bayesian hierarchical modeling framework. This can lead to additional insights on potentially modifiable risk factors that may contribute to elevated hospitalization and mortality risk and inform targeted patient care.

CHAPTER 2

Multivariate Varying Coefficient Spatiotemporal Model

2.1 Introduction

As of 2020, the number of individuals with prevalent end-stage kidney disease (ESKD) in the U.S. was $807,920$, where about 70% of patients with ESKD were on the life-sustaining treatment of dialysis (USRDS, 2022). Patients on dialysis have a high burden of complex comorbid conditions, and the overall mortality rates are 10-20 times higher than the general population (Bello et al., 2022; USRDS, 2022). Moreover, they experience frequent hospitalizations (around 2 per person-year (PPY)), which is a major source of morbidity and mortality (USRDS, 2022). Hence, modeling of the correlated multivariate outcome of hospitalization and survival rates in the U.S. dialysis cohort is of interest for more targeted patient monitoring.

Prior studies have shown important spatiotemporal variations in both hospitalization and survival rates. ESKD patients typically remain on dialysis for the duration of their lives or until kidney transplantation and their needs change as they stay longer on dialysis. Hence an important time index along which rates have been shown to vary is time since initiation of dialysis, where elevated hospitalization and mortality rates have been reported within the first 3 to 6 months of initiation of dialysis (Hickson et al., 2018; Estes et al., 2018; Li et al., 2018; Estes et al., 2016; Noordzij and Jager, 2014; de Jager et al., 2009). In addition to time-dynamic variation, there has been significant spatial variation reported in hospitalization and mortality rates across U.S., with highest rates observed in health service areas (HSAs) from the North East and lowest in HSAs from the West (Li et al., 2021, 2022; Erickson et al., 2019; USRDS, 2022). Hence, there is a compelling need for multivariate

modeling of hospitalization and mortality rates to identify time-dynamic/time-varying effects of region-specific risk factors (e.g., urbanicity and area deprivation index) that contribute to differences in longitudinal hospitalization and mortality risks observed across the U.S., while adjusting for spatiotemporal correlations in the data.

Multivariate varying coefficient modeling is an effective tool in studying time-varying effects of covariates on multivariate longitudinal response. Estimation typically relies on effective decompositions of the varying coefficient functions (VCFs) via local polynomial (Zhu et al., 2012; Kürüm et al., 2016; Zhang et al., 2021; Cai et al., 2007), or basis expansions (He et al., 2018; Yee and Wild, 1996; Wild and Yee, 1996; Yee and Mackenzie, 2002; Guo et al., 2022), similar to the univariate case. Literature on multivariate varying coefficient models that include space-time varying coefficients or incorporate spatiotemporal correlations in the data are limited. Gelfand et al. (2005) considered multivariate space-time varying coefficient models, while Congdon (2004) and Cheng et al. (2017) incorporated spatial correlations in the residuals. However, all three approaches model time as discrete. In our application of multivariate modeling of hospitalization and mortality risk across HSAs in U.S., space is modeled as discrete, since HSAs (regions with relatively self-contained infrastructure for the provision of hospital care) across U.S. are fixed and time is modeled as continuous, since we are interested in drawing inference along the continuous time index of duration on dialysis.

While there is extensive literature on multivariate spatiotemporal modeling where space and time are both modeled as discrete, the literature on discrete space and continuous time modeling is limited. For multivariate spatiotemporal models with discrete space and continuous time covariances, Zhang et al. (2006) proposed a separable structure for the multivariate space-time covariance, while Hepler et al. (2021) and Baer et al. (2021) considered nonseparable space-time covariances with quite restricted linear or user-defined parametric forms in the temporal modeling. Recently, Qian et al. (2024) proposed a multivariate spatiotemporal functional principal analysis model (MST-FPCA) with nonseparable space-time covariances, which couples dimension reduction in modeling nonparametric time trends through data-driven lower dimensional multivariate eigenfunctions with modeling of spatial correlations via a parametric conditional autoregressive (CAR) correlation structure. However, all these

works do not include time-varying effects of covariates.

To quantify the time-varying effects of risk factors on the correlated multivariate outcome of hospitalization and mortality rates in the dialysis cohort while incorporating spatiotemporal correlations in the data, where space is taken to be discrete and time is modeled continuous, we propose a novel multivariate varying coefficient spatiotemporal model (MV-VCSTM). In addition to modeling time-varying effects of covariates on the multivariate outcome, MV-VCSTM also includes region-specific time-varying multivariate random deviations in the model to account for the remaining multivariate spatiotemporal correlations in the data that are not explained by covariates. Similar to developments in our proposed MST-FPCA model (Qian et al., 2024), time-varying trends of the random deviations are modeled nonparametrically based on data-driven lower dimensional multivariate eigenfunction bases. Spatial correlations are induced through a CAR model assumed among region-specific scores, leading to a nonseparable structure on the multivariate space-time covariance. Estimation and inference in the proposed MV-VCSTM, simultaneously targeting multivariate time-varying coefficient functions and a multivariate spatiotemporal process poses a major computational challenge, especially in higher dimensional data applications. To achieve computational efficiency, the proposed estimation and inference relies on multiple computational savings. First, while the multivariate varying coefficient functions are expanded using thin-plate spline bases, region-specific random deviations are expanded via data-driven multivariate eigenfunction bases, which significantly reduces the number of parameters estimated associated with basis expansions. The multivariate eigenfunctions used in the expansion of the region-specific random deviations are derived based on only univariate functional principal component analysis (FPCA), avoiding smoothing over higher than 2-dimensional covariance surfaces, leading to major computational savings as the dimension of the multivariate outcome increases. The estimation of the multivariate eigenfunctions relies on univariate eigenfunctions and the estimated spatial correlation structure among the region-specific scores obtained from univariate FPCA. Estimation of this spatial correlation structure requires multivariate CAR (MCAR) modeling, which also poses challenges in higher dimensions. To handle this challenge, the MCAR structure is modeled via a linear

combination of independent latent Gaussian processes with a CAR correlation structure, effectively lowering the dimension in estimation. Once the multivariate eigenfunctions for bases expansions of the region-specific random deviations are obtained, we target the multivariate time-varying coefficient functions and region-specific deviation scores as well as the needed spatial parameters via Markov Chain Monte Carlo (MCMC) in a Bayesian hierarchical modeling framework.

The chapter is organized as follows. Section 2.2 introduces the proposed MV-VCSTM, and outlines the proposed estimation and inference procedures. The application to modeling of hospitalization and mortality rates in the U.S. dialysis cohort is based on data from the national USRDS database, containing longitudinal hospitalization and mortality rates from 367 HSAs across U.S. and is outlined in Section 2.3. Simulations to study the finite sample performance of the proposed methodology, including comparisons with multivariate varying coefficient models that do not incorporate spatiotemporal dependencies observed in the residuals are presented in Section 2.4, followed by a discussion given in Section 2.5.

2.2 Proposed Multivariate Varying Coefficient Spatiotemporal Model (MV-VCSTM)

2.2.1 Model Specification

Let $i = 1, \ldots, n$ index regions, $k = 1, \ldots, T$ index time (months) after transition to dialysis and $j = 1, \ldots, J$ ($J \geq 2$) index the dimensions of the J-dimensional multivariate outcome vector, $\boldsymbol{Y}_i(t_k) = \{Y_i^{(1)}(t_k), \ldots, Y_i^{(J)}(t_k)\}^\top$. In our data application, the multivariate outcome considered contains region-specific monthly hospitalization and mortality rates (i.e. $J = 2$). The region-specific rates are obtained as averages of dialysis facility-specific rates within regions (i.e., HSAs). Rates at the dialysis facility level are defined as the ratio of the total number of patient hospitalizations or deaths to the total patient follow-up time for that specific facility at month k. Even though they are obtained monthly, rates are multiplied by 12 to allow interpretations as per person-year (PPY), consistent with annual national

reporting of USRDS. Finally, region-specific rates are analyzed for a total of 24 months (i.e., 2 year follow-up) after transitioning to dialysis. Note that we opt to model the multivariate response as continuous data similar to previous works (Quick et al., 2013; Li et al., 2021; Qian et al., 2024) and is amenable for functional data modeling.

To study the time-varying effects of risk factors on the multivariate response (of hospitalization and mortality rates), the proposed MV-VCSTM includes P covariates $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{iP})^\mathsf{T}$ and P time-varying coefficient functions $\boldsymbol{\beta}^{(j)}(t) = \{\beta_1^{(j)}(t), \ldots, \beta_P^{(j)}(t)\}^\mathsf{T}$ associated with each covariate,

$$Y_i^{(j)}(t) = \boldsymbol{X}_i^\mathsf{T}\boldsymbol{\beta}^{(j)}(t) + U_i^{(j)}(t) + \epsilon_i^{(j)}(t), \quad \text{for} \quad j = 1, \ldots, J. \tag{2.1}$$

In (2.1), let $U_i^{(j)}(t)$ denote region-specific deviations, which are unobserved latent processes to capture the remaining spatiotemporal variations in the data after adjusting for time-varying effects of risk factors, while $\boldsymbol{\epsilon}_i(t) = \{\epsilon_i^{(1)}(t), \ldots, \epsilon_i^{(J)}(t)\}^\mathsf{T}$ denotes the i.i.d. measurement error with different variances along different dimensions of the multivariate outcome ($\epsilon_i^{(j)}(t) \sim_{ind} N(0, \sigma_j^2)$). Note that in the formulation in (2.1), the covariates are taken to be time-invariant, to mimic our data application. However, the developments can be extended easily to include time-varying covariates by changes to the design matrix, and the estimation algorithm will keep the same.

The region-specific deviations $\boldsymbol{U}_i(t) = \{U_i^{(1)}(t), \ldots, U_i^{(J)}(t)\}^\mathsf{T}$ are further decomposed via

$$\boldsymbol{U}_i(t) = \sum_{\ell=1}^{\infty} \rho_{i\ell}\boldsymbol{\psi}_\ell(t),$$

where $\rho_{i\ell}$ denotes the region-specific principal component (PC) scores, $\boldsymbol{\psi} = \{\psi^{(1)}, \ldots, \psi^{(J)}\}^\mathsf{T}$ denote functions, and $\boldsymbol{\psi}_\ell(t) = \{\psi_\ell^{(1)}(t), \ldots, \psi_\ell^{(J)}(t)\}^\mathsf{T}$ denote multivariate eigenfunctions which form an orthonormal system,

$$\langle\langle \boldsymbol{\psi}_\ell, \boldsymbol{\psi}_{\ell'} \rangle\rangle := \sum_{j=1}^{J} \langle \psi_\ell^{(j)}, \psi_{\ell'}^{(j)} \rangle_2 = \sum_{j=1}^{J} \int_{\mathcal{T}} \psi_\ell^{(j)}(t)\psi_{\ell'}^{(j)}(t)\mathrm{d}t = \delta_{\ell\ell'},$$

with $\delta_{\ell\ell'} = 1$ if $\ell = \ell'$, and $\delta_{\ell\ell'} = 0$ otherwise. Under the classic multivariate FPCA framework (Ramsay and Silverman, 2005; Happ and Greven, 2018), the multivariate PC scores $\{\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T} : \ell = 1, 2, \ldots\}$ are assumed to be uncorrelated, with zero means and $\text{Var}(\boldsymbol{\rho}_{i\ell}) = \lambda_\ell$, where λ_ℓ's denote the eigenvalues. Similar to univariate FPCA, the multivariate eigenfunctions describe directions of leading modes of variation in the different dimensions of the multivariate functional response, while the eigenvalues quantify the amount of variation explained along the identified modes of variation. In most applications the expansion can be truncated to include only a few number (denoted below with L) of eigencomponents explaining most of the variation in the data, leading to effective dimension reduction of the high-dimensional data.

To capture the remaining spatial correlation among regions, that is not explained by the time varying effects of regional risk factors, we induce a CAR structure on the region-specific PC scores $\rho_{i\ell}$, deviating from the classic FPCA framework. More specifically, let the $n \times n$ adjacency matrix $W = \{w_{ii'}\}$ describe the neighborhood structure of the regions, where $w_{ii'} = 1$ if regions i and i' $(i \neq i')$ are neighbors, denoted by $i \sim i'$, and $w_{ii'} = 0$ otherwise. By convention, the diagonal elements of W are set to zero. Further, let D denote the diagonal matrix with elements $d_i = \sum_{i \sim i'} w_{ii'}$, representing the total number of neighbors of regions. Then a Markov Random Field (Banerjee et al., 2015) specifies the full conditional distribution for the ℓth PC score $\rho_{i\ell}$ for region i as a weighted average of the ℓth PC scores from neighboring regions: $\rho_{i\ell} \mid \{\rho_{i'\ell}\}_{i \neq i'} \sim N(\nu \sum_{i' \sim i} w_{ii'} \rho_{i'\ell}/d_i, \alpha_\ell/d_i)$, where α_ℓ denotes the variance component and ν the spatial correlation parameter. Through Brook's lemma, the joint distribution of the ℓth PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T}$ takes the form $\boldsymbol{\rho}_\ell \sim N(\mathbf{0}, \alpha_\ell(D - \nu W)^{-1})$, where the scalar ν is constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues of the matrix $D^{-1/2} W D^{-1/2}$ for the precision matrix $(D - \nu W)/\alpha_\ell$ to be positive definite (Banerjee et al., 2015). The assumed CAR model facilitates borrowing of spatial information across neighboring regions, acting similar to a spatial smoother.

Note that even though different dimensions of $\boldsymbol{U}_i(t)$ share the same PC scores in the proposed modeling, MV-VCSTM is able to induce different spatiotemporal correlations across

different dimensions of $\boldsymbol{U}_i(t)$. This is because PC scores from different eigencomponents are multiplied by multivariate eigenfunctions which are allowed to vary across different dimensions of $\boldsymbol{U}_i(t)$. Moreover, the covariance of $\boldsymbol{U}_i(t)$ is a weighted linear combination (weighted by α_ℓ) of the spatial correlation induced across regions through the adjacency (W) and neighborhood matrices (D) multiplying different eigenfunctions, such that the imposed model on $\boldsymbol{U}_i(t)$ with a CAR correlation structure on the multivariate PC scores implies a nonseparable spatiotemporal correlation on $\boldsymbol{U}_i(t)$, which adds more flexibility to MV-VCSTM.

Note also that even though the measurement error variance σ_j^2 in (2.1) is assumed constant across time, it does't mean that the proposed MV-VCSTM assumes constant variance of the outcome across time. In the proposed modeling, after accounting for the time-varying effects of covariates on the multivariate outcome, the remaining spatiotemporal variation is modeled by the latent spatiotemporal processes $\boldsymbol{U}_i(t)$ via multivariate FPCA, which can model the leading directions of variation across time through multivariate eigenfunctions. Since the remaining temporal variation has already been accounted for through $\boldsymbol{U}_i(t)$, the measurement error variance is assumed to be constant across time to make the error components identifiable, which is a common assumption in FPCA (Ramsay and Silverman, 2005; Happ and Greven, 2018).

2.2.2 Estimation and Inference

The proposed algorithm begins by fitting an multivariate varying coefficient model to the observed data under the working independence assumption (Step 1), where we obtain the multivariate residuals, denoted by $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\intercal$, and use them as a proxy of the remaining spatiotemporal deviations $\boldsymbol{U_i}(t)$ in the data. Next, the multivariate eigenfunctions used in expansion of $\boldsymbol{U_i}(t)$ are targeted using univariate eigenfunctions from the different dimensions of the multivariate residual $\boldsymbol{E}_i(t)$ and the between eigencomponent dependencies of the univariate PC scores $\widehat{\boldsymbol{\xi}}$ of $\boldsymbol{E}_i(t)$ (reflecting the dependency structure between the different dimensions of the multivariate outcome) (Steps 2-4). Note that the multivariate conditional autoregressive (MCAR) modeling needed for targeting the between

eigencomponents covariance of $\widehat{\boldsymbol{\xi}}$ is fitted via a linear combination of independent latent Gaussian processes with a CAR correlation structure, achieving further dimension reduction and computational savings (Step 3).

The regular multivariate FPCA involves smoothing over four or more dimensional covariance matrices, which can be quite cumbersome due to "curse of dimensionality". Instead, the proposed MV-VCSTM relies only on the univariate FPCA employed on the different dimensions of $\boldsymbol{E}_i(t)$, which involves only smoothing over two-dimensional covariance matrices. By utilizing the dependency between the eigencomponents of the univariate response (through Σ_b) in deriving the multivariate eigenfunctions, the proposed modeling can avoid smoothing over higher dimensional covariance processes and the gain in computational efficiency becomes larger for larger J. Besides, the expansion of $\boldsymbol{U_i}(t)$ on data-driven multivariate eigenfunction bases, instead of a pre-determined basis set, can significantly lower the number of parameters targeted in the hierarchical model utilized in Step 5, in the sense that in most applications a small number (usually two to three) of multivariate eigenfunctions are enough to capture most of the variation in the data. In addition, and perhaps more importantly, the multivariate eigenfunctions provide extra interpretations in model building. Because they are estimated from the data, they represent dominant modes of variation in time.

Once the data-driven multivariate eigenfunction bases for $\boldsymbol{U_i}(t)$ are estimated and the multivariate varying coefficient functions (VCFs) $\boldsymbol{\beta}^{(j)}(t)$ are expanded onto thin-plate spline bases, spline coefficients needed for estimation of the multivariate VCFs, multivariate PC scores $\rho_{i\ell}$ and variance components (α_ℓ, ν and σ^2) are targeted via a Bayesian hierarchical model (Step 5), followed by inference based on posterior distributions of the penalized spline coefficients and the multivariate PC scores. While the proposed estimation procedure is outlined in the algorithm below, key features in each step are detailed further in this section. The R codes and a tutorial for implementing the MV-VCSTM algorithm are made publicly available on Github (https://github.com/qqian-biostat/MV-VCSTM).

Step 1: A multivariate varying-coefficient model is fit to the observed data under the working

Estimation Algorithm

Step 1: Fit a multivariate varying coefficient model to the observed data under the working independence assumption and obtain the multivariate residuals $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\mathsf{T}$.

Step 2: Employ univariate FPCA on $E_i^{(j)}(t)$ to target univariate eigenfunctions and the univariate PC score vector $\widehat{\boldsymbol{\xi}}$.

Step 3: Decompose dependencies among $\widehat{\boldsymbol{\xi}}$ into a between eigencomponent Σ_b and a within eigencomponent Σ_w (between region) variation via an MCAR correlation structure.

Step 4: Target the multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t)$ using the estimated univariate eigenfunctions and eigenvectors of the estimated Σ_b.

Step 5: Expanding the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions and the VCFs $\boldsymbol{\beta}^{(j)}(t)$ on thin-plate spline bases, target the VCFs $\boldsymbol{\beta}^{(j)}(t)$, multivariate PC scores ($\rho_{i\ell}$) and the variance components (α_ℓ, ν, and σ^2) within a Bayesian hierarchical modeling framework via MCMC.

independence assumption (referred to as MV-VCM): $Y_i^{(j)}(t) = \boldsymbol{X}_i^\mathsf{T}\boldsymbol{\beta}^{(j)}(t) + \epsilon_i^{(j)}(t)$, where $\epsilon_i^{(j)}(t)$ is assumed to be i.i.d. MV-VCM is implemented using the built-in `gam()` function in the R package `mgcv`, where thin-plate splines are used in estimation of the VCFs and the smoothing parameters are selected using restricted maximum likelihood (REML). The multivariate residuals from this initial fit, denoted by $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\mathsf{T}$, capture the remaining spatiotemporal variation in the data that is not explained by the time varying effects of risk factors and are used as a proxy for the region-specific deviations $\boldsymbol{U_i}(t)$.

Step 2: Univariate FPCA is employed on $E_i^{(j)}(t)$ separately along each dimension of the multivariate residuals. More specifically, in each dimension, we calculate the empirical covariance as $\widehat{G}^{(j)}(t,t') = \sum_{i=1}^n E_i^{(j)}(t)E_i^{(j)}(t')/n$. The diagonal entries of $\widehat{G}^{(j)}(t,t')$ are removed since they carry an additional error variance term, i.e., only off-diagonal elements are included as input data for the covariance surface smoothing step, where the bivariate penalized spline smoothers are applied via the built-in `gam()` function in the R package `mgcv`, yielding the univariate eigenfunction $\{\widehat{\phi}_m^{(j)}(t) : m = 1, \ldots, M_j\}$ and PC score estimates $\{\widehat{\xi}_{im}^{(j)} : i = 1, \ldots, n; m = 1, \ldots, M_j\}$. Smoothing parameters in the bivariate smoothing are selected by REML and the number of eigencomponents M_j are selected using the fraction of variance explained (FVE), where FVE $> 99\%$ is used in applications to retain enough information at this initial step of the algorithm.

Step 3: Next, the univariate PC score vector $\widehat{\boldsymbol{\xi}}_i^{\mathsf{T}} = (\widehat{\xi}_{i1}^{(1)}, \ldots, \widehat{\xi}_{iM_1}^{(1)}, \ldots, \widehat{\xi}_{i1}^{(J)}, \ldots, \widehat{\xi}_{iM_J}^{(J)})$ is modeled using an MCAR correlation structure. Let $\Xi = [\widehat{\boldsymbol{\xi}}_1, \ldots, \widehat{\boldsymbol{\xi}}_n]^{\mathsf{T}}$ denote the $n \times M^+$ score matrix, combining the score vectors from all subjects, where $M^+ = M_1 + M_2 + \ldots + M_J$ denotes the total number of eigencomponents retained across all J dimensions of $\boldsymbol{E}_i(t)$. Stacking the columns of Ξ leads to the $nM^+ \times 1$ score vector $\widehat{\boldsymbol{\xi}} = \text{vec}(\Xi)$ that is modeled as a linear combination of M^+ independent latent Gaussian processes with a CAR correlation structure (Jin et al., 2007),

$$\underset{nM^+\times 1}{\widehat{\boldsymbol{\xi}}} = \underset{nM^+\times 1}{\boldsymbol{\xi}} + \underset{nM^+\times 1}{\mathbf{e}} = (\underset{M^+\times M^+}{A} \otimes \underset{n\times n}{I}) \underset{nM^+\times 1}{\mathbf{f}} + \underset{nM^+\times 1}{\mathbf{e}}. \tag{2.2}$$

In (2.2), $A = \{a_{\ell\ell'}\}$, $1 \leq \ell' \leq \ell \leq M^+$, denotes an $M^+ \times M^+$ full rank lower triangular matrix and $\mathbf{e} \sim N(\mathbf{0}, \tau^2 I_{nM^+})$ denotes the vector of measurement errors, assumed to be uncorrelated with $\boldsymbol{\xi}$. Since each $n \times 1$ latent Gaussian process has a CAR correlation structure, i.e., $\mathbf{f}_\ell \sim N(\mathbf{0}, (D - \nu W)^{-1})$, $\ell = 1, \ldots, M^+$, where D and W are as defined in Section 2.2.1 and ν denotes the common spatial smoothing parameter, the joint distribution of the vector of stacked latent Gaussian processes $\mathbf{f} = (\mathbf{f}_1^{\mathsf{T}}, \ldots, \mathbf{f}_{M^+}^{\mathsf{T}})^{\mathsf{T}}$ follows $N(\mathbf{0}, I_{M^+} \otimes (D - \nu W)^{-1})$. This leads to the decomposition of the covariance Σ of $\boldsymbol{\xi}$ into between eigencomponent ($\Sigma_b \equiv AA^{\mathsf{T}}$) and within eigencomponent ($\Sigma_w \equiv (D - \nu W)^{-1}$) variation:

$$\begin{aligned}
\underset{nM^+\times nM^+}{\Sigma} &= (\underset{M^+\times M^+}{A} \otimes \underset{n\times n}{I}) \text{Bdiag}\{\underset{n\times n}{(D - \nu W)^{-1}}, \ldots, (D - \nu W)^{-1}\}(A^{\mathsf{T}} \otimes I) \\
&= \underset{M^+\times M^+}{AA^{\mathsf{T}}} \otimes \underset{n\times n}{(D - \nu W)^{-1}} \equiv \underset{M^+\times M^+}{\Sigma_b} \otimes \underset{n\times n}{\Sigma_w},
\end{aligned}$$

where Bdiag$(\cdot)$ denotes a block-diagonal matrix and $\boldsymbol{\xi}$ follows MCAR (ν, Σ_b). While the between eigencomponent variation captures the dependency between the eigencomponents from univariate expansions, the within eigencomponent (between region) variation captures the spatial dependency among the regional units. Note that representation of $\boldsymbol{\xi}$ as a linear combination of independent spatial Gaussian processes modeled with a CAR correlation structure, via the lower triangle matrix A, allows for easy implementation of the MCAR model in WinBUGS, without the need for operations with large covariance matrices and

without the need to check for positive-definiteness of covariance matrices in each iteration of the algorithm.

The parameters of the MCAR model are targeted via MCMC. Elementwise priors are imposed on the lower-triangular matrix A: $a_{\ell\ell} \sim$ Log-normal$(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}})$ and $a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$ for $1 \leq \ell' < \ell \leq M^+$. In addition, an Inverse Gamma (IG) (a_{τ^2}, b_{τ^2}) prior is imposed on the measurement error variance τ^2, and a Uniform prior is used for the spatial correlation parameter ν with the parameters constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues of the matrix $D^{-1/2}WD^{-1/2}$ (denoted by a_ν, b_ν, respectively). The posterior distribution we seek can be expressed as

$$\begin{aligned}\pi(a_{\ell\ell}, a_{\ell\ell'}, \nu, \tau^2 \mid \widehat{\boldsymbol{\xi}}) \propto\ & N(\widehat{\boldsymbol{\xi}} \mid \boldsymbol{\xi}, \tau^2) \times N\left(\boldsymbol{\xi} \mid \mathbf{0}, AA^{\mathsf{T}} \otimes (D - \nu W)^{-1}\right) \\ & \times \prod_{\ell=1}^{M^+} \text{Log-normal}(a_{\ell\ell} \mid \mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}}) \times \prod_{\ell=1}^{M^+} \prod_{\ell'=1}^{\ell-1} N(a_{\ell\ell'} \mid \mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}}) \\ & \times \text{Unif}(\nu \mid a_\nu, b_\nu) \times \text{IG}(\tau^2 \mid a_{\tau^2}, b_{\tau^2}).\end{aligned}$$

Model parameters are sampled from the posterior distributions using MCMC with Gibbs sampling and random walk metropolis as implemented in WinBUGS.

Step 4: The multivariate eigenfunctions are targeted using the univariate eigenfunctions $\widehat{\phi}_m^{(j)}(t)$ of $E_i^{(j)}(t)$ estimated in Step 2 and the eigenvectors $\widehat{\mathbf{c}}_\ell$, $\ell = 1, \ldots, M^+$, of the between eigencomponent variation $\widehat{\Sigma}_b$ targeted in Step 3 as

$$\widehat{\psi}_\ell^{(j)}(t) = \sum_{m=1}^{M_j} [\widehat{\mathbf{c}}_\ell]_m^{(j)} \widehat{\phi}_m^{(j)}(t), \quad \ell = 1, \ldots, M^+, \quad j = 1, \ldots, J. \tag{2.3}$$

The between eigencomponent dependence is incorporated in estimation of the multivariate eigenfunctions in (2.3) via the linear combination of the univariate eigenfunctions with weights equal to the mth entry of $[\widehat{\mathbf{c}}_\ell]^{(j)} \in \mathbb{R}^{M_j}$, the jth block of the ℓth eigenvector $\widehat{\mathbf{c}}_\ell$ of $\widehat{\Sigma}_b$ (equal to the posterior mean of Σ_b). For justification of (2.3), we refer the readers to Appendix A.1, where we argue for (2.3) in the context of a spatiotemporal model via the link we establish between the variance components of the response and the eigenvalues

of the between eigencomponent covariance of the univariate PC scores. The form in (2.3) can follow from the fact that the variance components α_ℓ's are also the eigenvalues of the between eigencomponent covariance matrix Σ_b, which links the multivariate and univariate eigenfunctions with the eigenvectors of Σ_b.

Step 5: Expanding the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions from Step 4 and the VCFs $\boldsymbol{\beta}^{(j)}(t)$ on thin-plate spline bases, the VCFs, multivariate PC scores ($\rho_{i\ell}$) and the variance components (α_ℓ, ν, and σ^2) are targeted via a hierarchical Bayesian modeling framework.

Consider the expansion of the VCFs $\boldsymbol{\beta}^{(j)}(t) = \{\beta_1^{(j)}(t), \ldots, \beta_P^{(j)}(t)\}^\mathsf{T}$ onto penalized low-rank thin-plate spline bases:

$$\beta_p^{(j)}(t) \approx u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} v_{pq}^{(j)} |t - \kappa_q|^3, \quad \text{for} \quad p = 1, \ldots, P,$$

where $\kappa_1 < \kappa_2 < \ldots < \kappa_Q$ denote the fixed knots and $u_{p0}^{(j)}$, $u_{p1}^{(j)}$ and $\boldsymbol{v}_p^{(j)} = (v_{p1}^{(j)}, \ldots, v_{pQ}^{(j)})^\mathsf{T}$ denote the expansion coefficients. The number of knots Q utilized in applications is typically selected between $3 - 20$ to ensure the desired flexibility (Ruppert et al., 2003; Crainiceanu et al., 2005), and κ_q is taken to be the sample quantile of the time points corresponding to probability $q/(Q+1)$.

The P-splines, including a penalty matrix Ω_Q with (q', q)th entry $|\kappa_{q'} - \kappa_q|^3$, penalizing the coefficients of $|t - \kappa_q|^3$, can be fitted using the transformed model

$$\beta_p^{(j)}(t_k) \approx u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}, \quad \text{for} \quad p = 1, \ldots, P,$$

where $\widetilde{\boldsymbol{v}}_p^{(j)} \equiv (\widetilde{v}_{p1}^{(j)}, \ldots, \widetilde{v}_{pQ}^{(j)})^\mathsf{T} = \Omega_Q^{1/2} \boldsymbol{v}_p^{(j)}$ with $\text{cov}(\widetilde{\boldsymbol{v}}_p^{(j)}) = \sigma^2_{\widetilde{v}_{pj}} I_Q$ and z_{kq} denotes the (k, q)th entry of the transformation $Z = Z_Q \Omega_Q^{-1/2}$ of the design matrix Z_Q with kth row equaling $\{|t_k - \kappa_1|^3, \ldots, |t_k - \kappa_Q|^3\}$ (Crainiceanu et al., 2005).

The hierarchical model combining expansions of the VCFs on thin-plate spline bases and

expansion of $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions $\widehat{\psi}_\ell^{(j)}(t)$ can be given as

$$Y_i^{(j)}(t_k) = \sum_{p=1}^{P} X_{ip}\left(u_{p0}^{(j)} + u_{p1}^{(j)}t + \sum_{q=1}^{Q}\widetilde{v}_{pq}^{(j)}z_{kq}\right) + \sum_{\ell=1}^{M^+}\rho_{i\ell}\widehat{\psi}_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k),$$

for $j = 1, \ldots, J$ and $k = 1, \ldots, T$, where $\widetilde{v}_{pq}^{(j)} \sim N(0, \sigma^2_{\widetilde{v}_{pj}})$, $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T} \sim N(\mathbf{0}, \alpha_\ell(D - \nu W)^{-1})$ and $\epsilon_i^{(j)}(t_k) \sim N(0, \sigma_j^2)$. Using normal priors for $u_{p0}^{(j)} \sim N(0, \sigma^2_{u_{p0j}})$ and $u_{p1}^{(j)} \sim N(0, \sigma^2_{u_{p1j}})$, inverse Gamma priors for the variance components $\sigma^2_{\widetilde{v}_{pj}} \sim \text{IG}(a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}})$, $\sigma_j^2 \sim \text{IG}(a_{\sigma_j^2}, b_{\sigma_j^2})$, $\alpha_\ell \sim \text{IG}(a_{\alpha_\ell}, b_{\alpha_\ell})$ and a Uniform prior for $\nu \sim \text{Unif}(a_\nu, b_\nu)$, the posterior distribution we seek can then be expressed as

$$\begin{aligned}
&\pi(u_{p0}^{(j)}, u_{p1}^{(j)}, \widetilde{v}_p^{(j)}, \boldsymbol{\rho}_\ell, \nu, \alpha_\ell, \sigma_j^2 \mid Y_i^{(j)}(t), \boldsymbol{X}_i, Z, \widehat{\psi}_\ell^{(j)}(t), M^+) \\
\propto\ & \prod_{i=1}^{n}\prod_{k=1}^{T}\prod_{j=1}^{J} N\left(\boldsymbol{Y}_i^{(j)}(t_k) \mid \{\sum_{p=1}^{P} X_{ip}(u_{p0}^{(j)} + u_{p1}^{(j)}t + \sum_{q=1}^{Q}\widetilde{v}_{pq}^{(j)}z_{kq}) + \sum_{\ell=1}^{M^+}\rho_{i\ell}\psi_\ell^{(j)}(t_k)\}, \sigma_j^2\right) \\
\times\ & \prod_{j=1}^{J}\text{IG}(\sigma_j^2 \mid a_{\sigma_j^2}, b_{\sigma_j^2}) \times \prod_{j=1}^{J}\prod_{p=1}^{P} N\left(u_{p0}^{(j)} \mid 0, \sigma^2_{u_{p0j}}\right) \times \prod_{j=1}^{J}\prod_{p=1}^{P} N\left(u_{p1j}^{(j)} \mid 0, \sigma^2_{u_{p1j}}\right) \\
\times\ & \prod_{j=1}^{J}\prod_{p=1}^{P}\prod_{q=1}^{Q} N\left(\widetilde{v}_{pq}^{(j)} \mid 0, \sigma^2_{\widetilde{v}_{pj}}\right) \times \prod_{j=1}^{J}\prod_{p=1}^{P}\prod_{q=1}^{Q}\text{IG}\left(\sigma^2_{\widetilde{v}_{pj}} \mid a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}}\right) \\
\times\ & \prod_{\ell=1}^{M^+} N\left(\boldsymbol{\rho}_\ell \mid \mathbf{0}, \alpha_\ell(\mathrm{D} - \nu\mathrm{W})^{-1}\right) \times \prod_{\ell=1}^{M^+}\text{IG}(\alpha_\ell \mid a_{\alpha_\ell}, b_{\alpha_\ell}) \times \text{Unif}(\nu \mid a_\nu, b_\nu).
\end{aligned}$$

Model parameters are sampled from the posterior distributions using MCMC as implemented in WinBUGS. Since M^+ is the total number of univariate eigencomponents retained in the FPCA decompositions of $E_i^{(j)}(t)$ across the J dimensions, the total number (L) of multivariate eigencomponents retained in expansions of $\boldsymbol{U}_i(t)$ could be chosen to be smaller than M^+ in applications where M^+ may be large. Following (Qian et al., 2024), we recommend the use of the estimated variance components $\widehat{\alpha}_\ell$, $\ell = 1, \ldots, M^+$ in formulating the fraction of variance explained (FVE) in the choice of L, since they are established as the eigenvalues of the between eigencomponent covariance of the univariate PC scores (see Appendix A.1).

VCFs $\beta_p^{(j)}(t)$ are targeted by mean of the posterior draws for $u_{p0}^{(j)} + u_{p1}^{(j)}t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}$, based on posterior draws for $u_{p0}^{(j)}$, $u_{p1}^{(j)}$ and $\widetilde{\boldsymbol{v}}_p^{(j)}$. Simultaneous credible bands for the VCFs are obtained along each dimension j as follows. Let $f(t)$ denote a single VCF observed at time points t_k for $k = 1, 2, \ldots, T$ in one dimension. Further let $\widehat{f}(t)$ and $\text{SD}\{f(t)\}$ denote the mean and standard deviation of $f(t)$ based on a total of B MCMC samples $f^{(b)}(t)$ and let c_δ denote the $(1-\delta)$ sample quantile of $\max_{k=1,\ldots,T} \mid \{f^{(b)}(t_k) - \widehat{f}(t_k)\}/\text{SD}\{f(t_k)\} \mid$, $b = 1, 2, \ldots, B$. Then the $(1-\delta)$ simultaneous credible band (CB) for $f(t)$ in that dimension is given by $[\widehat{f}(t_k) \pm c_\delta \text{SD}\{f(t_k)\}]$ (Crainiceanu et al., 2007). In addition, using the estimated multivariate eigenfunctions from Step 4 and taking $\widehat{\boldsymbol{\beta}}^{(j)}(t)$ and $\widehat{\rho}_{i\ell}$ as the mean of the posterior samples on the VCFs and the multivariate eigenscores, respectively, region-specific prediction of the multivariate outcome is targeted by

$$\widehat{Y}_i^{(j)}(t) = \boldsymbol{X}_i^{\mathsf{T}} \widehat{\boldsymbol{\beta}}^{(j)}(t) + \sum_{\ell=1}^{L} \widehat{\rho}_{i\ell} \widehat{\psi}_\ell^{(j)}(t_k) \quad \text{for} \quad j = 1, \ldots, J.$$

When different dimensions of the residuals $\boldsymbol{E}_i(t)$ obtained in Step 1 have large differences in their range or if they exhibit different amounts of variation, standardization of the residuals may be considered along each dimension $E_i^{(j)}(t)$, to obtain interpretable multivariate functional principal components (Jacques and Preda, 2014; Chiou et al., 2014; Happ and Greven, 2018). The standardization can be carried out by $\widetilde{E}^{(j)}(t) = s_j^{1/2} E^{(j)}(t)$ where $s_j = \left\{ \int_{\mathcal{T}} \widehat{\text{Var}} \left[E^{(j)}(t) \right] \text{d}t \right\}^{-1}$, to guarantee that the integrated variance along the rescaled residuals equals one (Happ and Greven, 2018). In this way all dimensions of the multivariate residuals contribute equal amounts of variation to the analysis, similar to the standardization in classical multivariate principal component analysis (PCA).

2.3 Modeling Hospitalization and Mortality Risk Among Patients on Dialysis

2.3.1 Description of the USRDS Study Cohort and Region-Specific Risk Factors

The USRDS is a large national database which collects data on nearly all patients with ESKD in the U.S. We include patients aged 18 years or older who transitioned to dialysis between January 1, 2005, and September 30, 2013 in our study cohort and include a maximum follow-up of two years (beginning on day 91 of dialysis, after 90 days to establish a stable treatment modality, and with the last date of follow-up on December 31, 2015). Regional units are taken to be Health Service Areas (HSAs) with relatively self-contained infrastructure for the provision of hospital care in the contiguous U.S., including the District of Columbia. Region-specific multivariate outcome is obtained as the average of facility-specific hospitalization and mortality rates, calculated monthly per person-year (PPY) over the two year follow-up. In obtaining facility-specific rates, we exclude facilities with less than 10 patients in any one month follow-up interval to obtain stable facility-specific hospitalization and mortality rates over time (17.4% facilities excluded). At the region level, we merge HSAs to guarantee that each region contains at least 5 facilities (reducing available HSAs from 719 to 402) and merge HSAs to ensure that the total number of months with zero mortality rate is limited to one-third of the two-year follow-up, to further stabilize region-specific inference. The final study cohort contains 367 regions/HSAs after the second merging, where majority (61.6%) of the resulting 367 regions are unmerged, and 15.0%, 8.4%, and 15.0% of the regions consist of merging 2, 3, and > 3 original HSAs, respectively. The mean region-specific hospitalization and mortality rate in the final cohort are 1.810 and 0.073 PPY, respectively. Detailed descriptions of the study cohort, exclusion rules and region merging algorithms are deferred to Appendix B.1.

The proposed modeling includes urbanicity, area deprivation index (ADI) and medical underservice index (MUI) as region-specific covariates. Urbanicity is assigned by categorizing HSAs into the three classes of large metropolitan, small metropolitan or non-

metropolitan regions. Categorization of HSAs are assigned according to the class that majority of the counties within each HSA fall into, determined by the urban-rural classification scheme of the National Center for Health Statistics (https://www.cdc.gov/nchs/data_access/urban_rural.htm). Non-metropolitan regions are taken as the reference group. ADI, consisting of 17 education, employment, housing-quality, and poverty measures, captures the socio-economic status of HSAs (Kind and Buckingham, 2018). The rank-based index, taking on values between 0 and 100 (higher values correspond to lower socio-economic status and higher deprivation), is available at the census block group level (available at https://www.neighborhoodatlas.medicine.wisc.edu), and is averaged over census block groups within each HSA to arrive at the HSA level index. The last covariate, MUI, takes on values between 0 and 1 to reflect the medical service availability within each HSA (higher indices corresponding to higher underservice). The index is released by the Health Resources and Services Administration at the census tract/county subdivision level at https://data.hrsa.gov/tools/shortage-area, where medically underserved tracts/subdivisions are areas with too few primary care providers, high infant mortality, high poverty or a high elderly population. The proportion of census tracts/county subdivisions that are designated as underserved is first targeted for each county, and county MUIs are then averaged within each HSA to arrive at the HSA-level MUI index. ADI and MUI are mean-centered for ease of interpretation in modeling.

2.3.2 Results

2.3.2.1 Time-Varying Effects of Region-Specific Covariates

The estimated time-varying effects of the region-specific covariates are given in Figure 2.1. The time-varying y-intercepts represent the hospitalization (Figure 2.1 (a)) and mortality (Figure 2.1 (b)) rates of an "average" non-metropolitan region (reference group) with mean ADI of 58.948 and MUI of 0.502. The highest rates in both outcomes (2.023 PPY for hospitalization and 0.083 PPY for mortality) are observed within the first three months after transitioning to dialysis where both rates steadily decline during the first year on dialysis.

While hospitalization rates continue to decline after the first year, mortality rates remain relatively stable in the second year on dialysis. Figures 2.1 (c)-(j) display the estimated effects of risk factors (solid) with simultaneous 95% credible bands (dashed), overlaying horizontal lines at zero (gray) included for reference. Large metropolitan regions have significantly higher hospitalization rates than non-metropolitan regions (reference group), especially at initiation of dialysis, where the effect gradually weakens in the second year on dialysis. Similarly, ADI is significantly positively associated with hospitalizations, especially during the first and a half years on dialysis, suggesting that regions with higher deprivation level have higher hospitalization risk. The effects of small metropolitan regions and MUI are not found significant on hospitalizations as the credible bands include zero, indicating that small metropolitan regions have similar hospitalization rates to non-metropolitan regions, and medical service availability is not found to be significantly associated with hospitalization rates. Note that none of the covariates have significant effects on mortality, indicating that hospitalization rates are more susceptible to urbanicity and socioeconomic status of a region than mortality in the dialysis population.

To further assess the effects of the identified risk factors (being a large metropolitan region with a high ADI index) on hospitalization and mortality rates, Figure B.1 in Appendix B displays predicted hospitalization and mortality trajectories for two hypothetical regions, one that is a large metropolitan with 95% percentile ADI, and the other a non-metropolitan region with 5% percentile ADI (while MUI kept at the mean level of 0.502). The 95% simultaneous credible bands (shaded) for the predicted hospitalization and mortality trajectories are formed based on the draws from $\widetilde{Y}_i^{(j)}(t) = \boldsymbol{X}_i^{\mathsf{T}} \widehat{\boldsymbol{\beta}}^{(j)}(t)$ without the region-specific random effects, corresponding to each posterior draw of the multivariate varying coefficient functions, and following the algorithm outlined at the end of Section 2.2.2 for forming 95% simultaneous credible bands with $f(t) = Y_i^{(j)}(t)$. A large metropolitan with 95% percentile ADI is predicted to have higher hospitalization and mortality rates, compared to a non-metropolitan region with 5% percentile ADI, although the difference is not found significant with credible bands overlapping throughout the two-year follow-up for both outcomes. Moreover, while the difference of hospitalization rates between the two regions is relatively stable across the

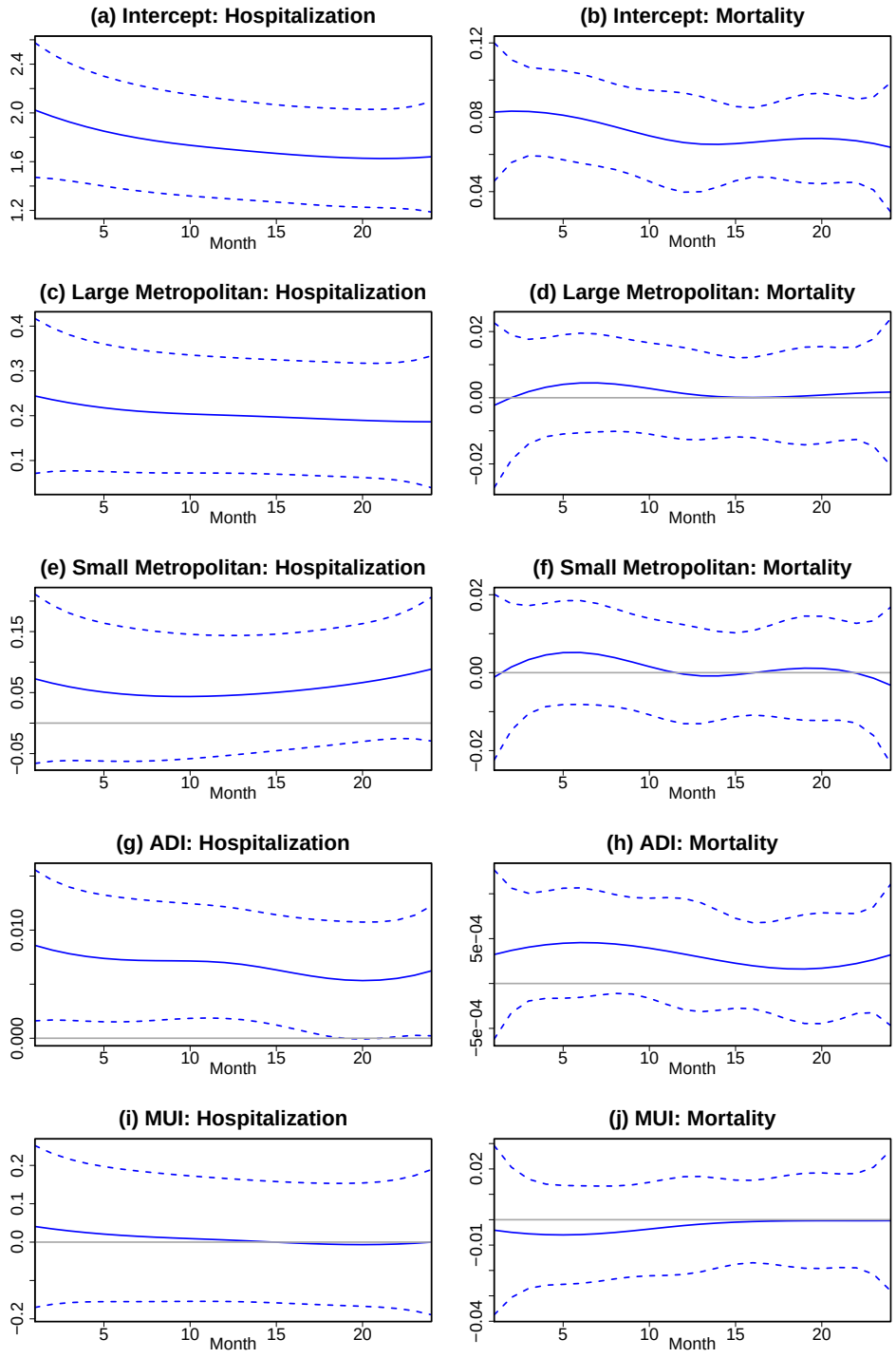

Figure 2.1: Estimated time-varying effects (solid) of covariates: y-intercept, large metropolitan, small metropolitan, area deprivation index (ADI) (centered) and medical underservice index (MUI) (centered), along with their 95% simultaneous credible bands (dashed), in modeling hospitalization (a, c, e, g, i) and mortality (b, d, f, h, j) rates among the U.S. dialysis population. Horizontal lines at zero are included in gray for reference.

two-year follow-up, the difference of morality rates between these two regions is larger during the first year on dialysis.

2.3.2.2 Spatiotemporal Patterns of Hospitalization and Mortality Rates

In addition to modeling the time-varying effects of the above risk factors, the proposed MV-VCSTM also models the remaining spatiotemporal patterns in the data via time-varying region-specific stochastic deviations $\boldsymbol{U}_i(t)$. Multivariate FPCA is utilized in reducing the dimension of $\boldsymbol{U}_i(t)$, where spatial correlations are induced among the region-specific eigenscores. A total of four and three eigencomponents are retained in univariate FPCA decompositions of hospitalization and mortality, respectively, explaining 99.02% and 99.17% of the total variation. The estimated multivariate eigenfunctions based on the univariate FPCA decompositions are given in Figure 2.2 with estimated spatial variance parameters $\widehat{\alpha}_1 = 6.964$, $\widehat{\alpha}_2 = 0.259$, $\widehat{\alpha}_3 = 0.151$, $\widehat{\alpha}_4 = 0.149$, $\widehat{\alpha}_5 = 0.021$, $\widehat{\alpha}_6 = 0.017$, and $\widehat{\alpha}_7 = 0.016$. The leading four multivariate eigenfunctions (given in Figures 2.2 (a)-(b)) mainly explain variation in hospitalization rates, while the last three (given in Figure 2.2 (c)-(d)) explain variation in mortality. Among the multivariate functions that mostly explain variability in hospitalization, the leading eigenfunction describes constant variation throughout the two-year follow-up, while the remaining three highlight variation in the first and last six months on dialysis (second eigenfunction), variation at 18 months and at end of the two year follow-up (third eigenfunction) and finally at one year and 18 months on dialysis (fourth eigenfunction). Among the multivariate eigenfunctions that mostly explain variation in mortality, the leading one (5th eigenfunction) explains constant variation over the two-year follow-up (similar to the leading eigenfunction explaining mostly constant variation in hospitalization) (given in Figure 2.2 (c)-(d)). The sixth multivariate eigenfunction explains variation in mortality at one year and 18 months on dialysis, while the seventh highlighting variation in mortality at 5 and 10 months on dialysis. The leading time-varying variation in hospitalization observed within the first six months of dialysis is consistent with higher hospitalization rates observed at initiation of dialysis, while higher variation in the last six months of follow-up may be related to the decrease in sample size of the study cohort at the end of the two year

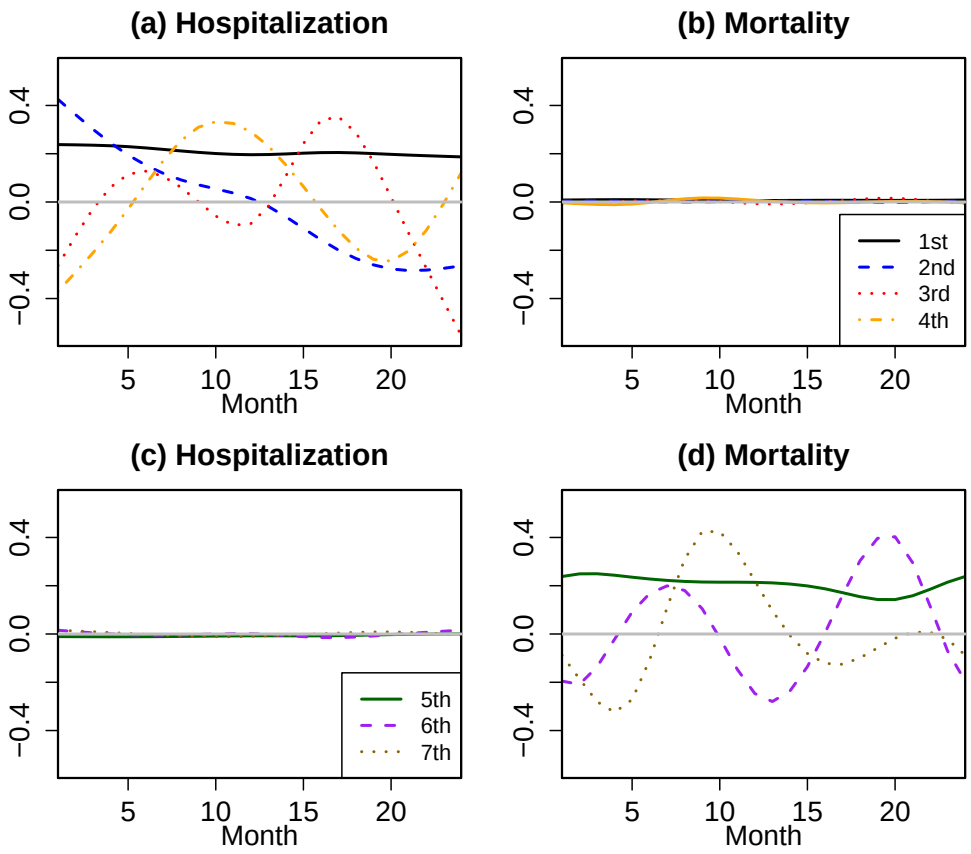

Figure 2.2: Estimated multivariate eigenfunctions for hospitalization (a, c) and mortality (b, d). The leading four multivariate eigenfunctions (a, b) mainly explain variation in hospitalization, while the last three multivariate eigenfunctions (c, d) mainly explain variation in mortality. The zero horizontal lines are included in gray for reference.

follow-up.

The estimated spatial correlation parameter of $\widehat{\nu} = 0.875$, inducing correlations between neighboring HSAs ranging from 0.19 to 0.60, confirm that there is still significant spatiotemporal variation remaining in the data, after adjusting for time-varying effects of covariates. To visualize the remaining spatial patterns in the data, Figure 2.3 (a)-(b) displays the residuals $\boldsymbol{E}_i(t)$ (averaged across time) obtained from the initial multivariate varying coefficient model (MV-VCM) fit to the data in Step 1 under the working independence assumption. Also displayed in Figure 2.3 (c)-(d) are the predicted region-specific deviations $\boldsymbol{U}_i(t)$ averaged across follow-up time. These maps confirm the significant spatial variation remaining in the data after adjusting for time-varying effects of covariates, i.e., the hospitalization and

mortality rates in the west regions tend to be overall lower compared with the rates in the east regions. In addition, the similarity between two groups of maps, i.e., maps of residuals from MV-VCM (Figure 2.3 (a)-(b)) and maps of predicted region-specific deviations $\boldsymbol{U}_i(t)$ (Figure 2.3 (c)-(d)) show that the predicted region-specific deviations are able to capture the remaining spatial variation in the data effectively. By comparison, the residuals from the proposed MV-VCSTM are quite small for hospitalizations, with larger deviations observed in a few HSAs in the harder to model outcome of mortality (Figure 2.3 (e)-(f)). There does not seem to be an obvious spatial correlation in the residuals from MV-VCSTM in both outcomes, implying that the proposed MV-VCSTM is able to model spatial correlations in the data effectively.

To assess the spatial and temporal variation jointly in hospitalization and mortality risk while accounting for the time-varying effects of region-specific covariates on the multivariate outcome, Figure 2.4 displays raw hospitalization (a) and mortality (b) rates as well as region-specific predictions from the full model at 3 month, 12 months and 18 months after initiation of dialysis. Overall both set of maps highlight elevated hospitalization and mortality rates in the "band" from Massachusetts to southern Texas (darker blue), and a decreasing trend in both rates for longer follow-up times on dialysis (consistent with the decreasing trends of the estimated y-intercepts for both outcomes). There are also outcome-specific trends with different spatial patterns among hospitalization and mortality risk in some regions, such as some HSAs in northern California, Oregon, Montana, and Idaho have relatively low hospitalization rates but elevated mortality rates, whereas some HSAs in Florida and Arizona with relatively high hospitalization rates have low mortality rates. Note that the predicted maps correspond closely to the raw maps, especially for hospitalization risk, indicating the satisfactory fit of the proposed MV-VCSTM model. Mortality risk is harder to predict than hospitalization risk as we observe from Figure 2.3 (e)-(f) and the prediction errors reported next, however, the proposed MV-VCSTM leads to the smallest prediction error also for mortality, compared to reduced models that ignore the spatiotemporal correlations in the region-specific deviations.

Next, the proposed MV-VCSTM fits are compared to fits from two reduced multivariate

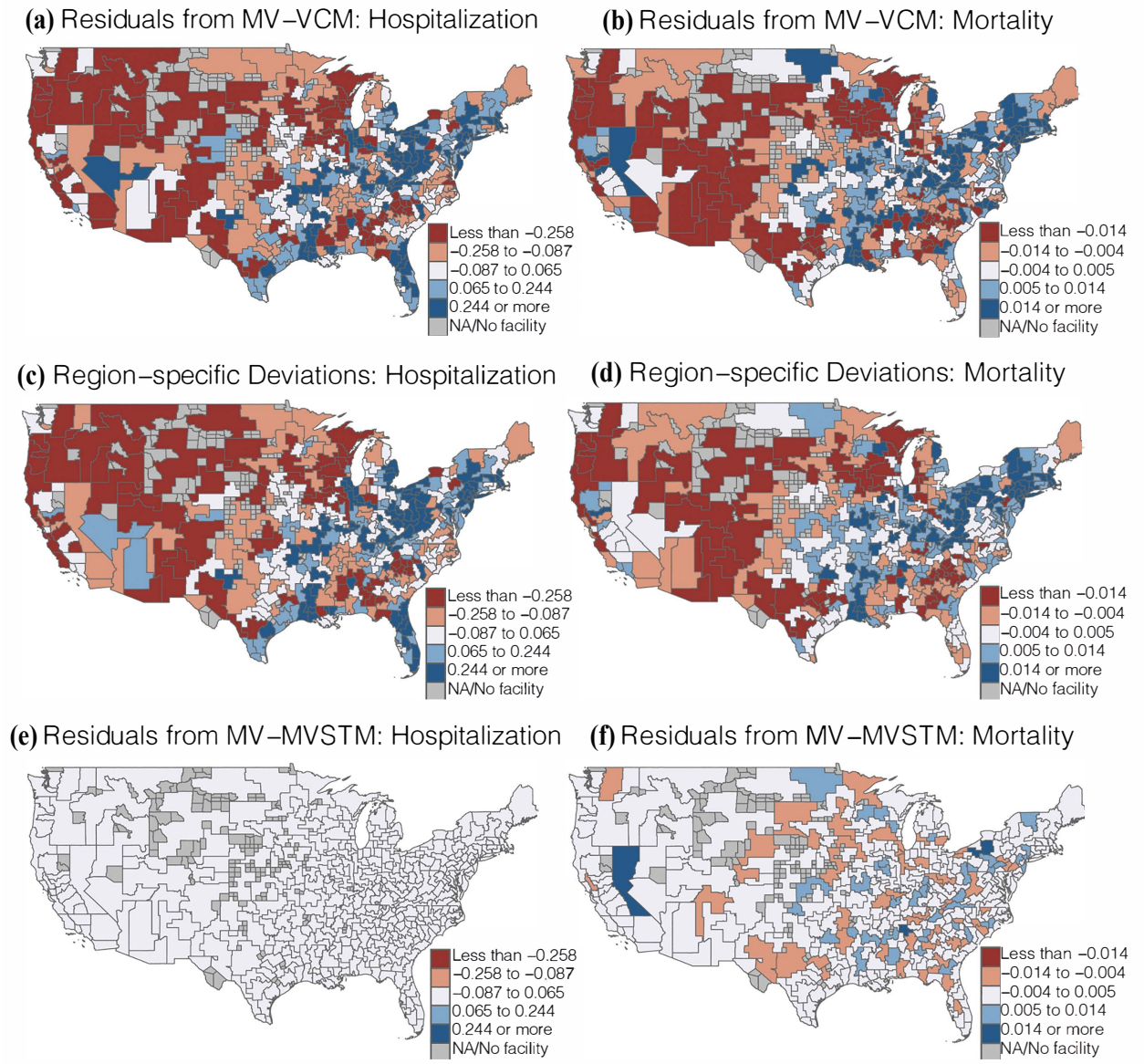

Figure 2.3: The residuals $\boldsymbol{E}_i(t)$ (averaged across time) obtained from the initial multivariate varying coefficient model fit to the data under the working independence assumption (a-b), the predicted region-specific deviations $\boldsymbol{U}_i(t)$ averaged across follow-up time (c-d), and the residuals (averaged across time) obtained from the proposed MV-VCSTM (e-f) for hospitalization (a, c, e) and mortality (b, d, f).

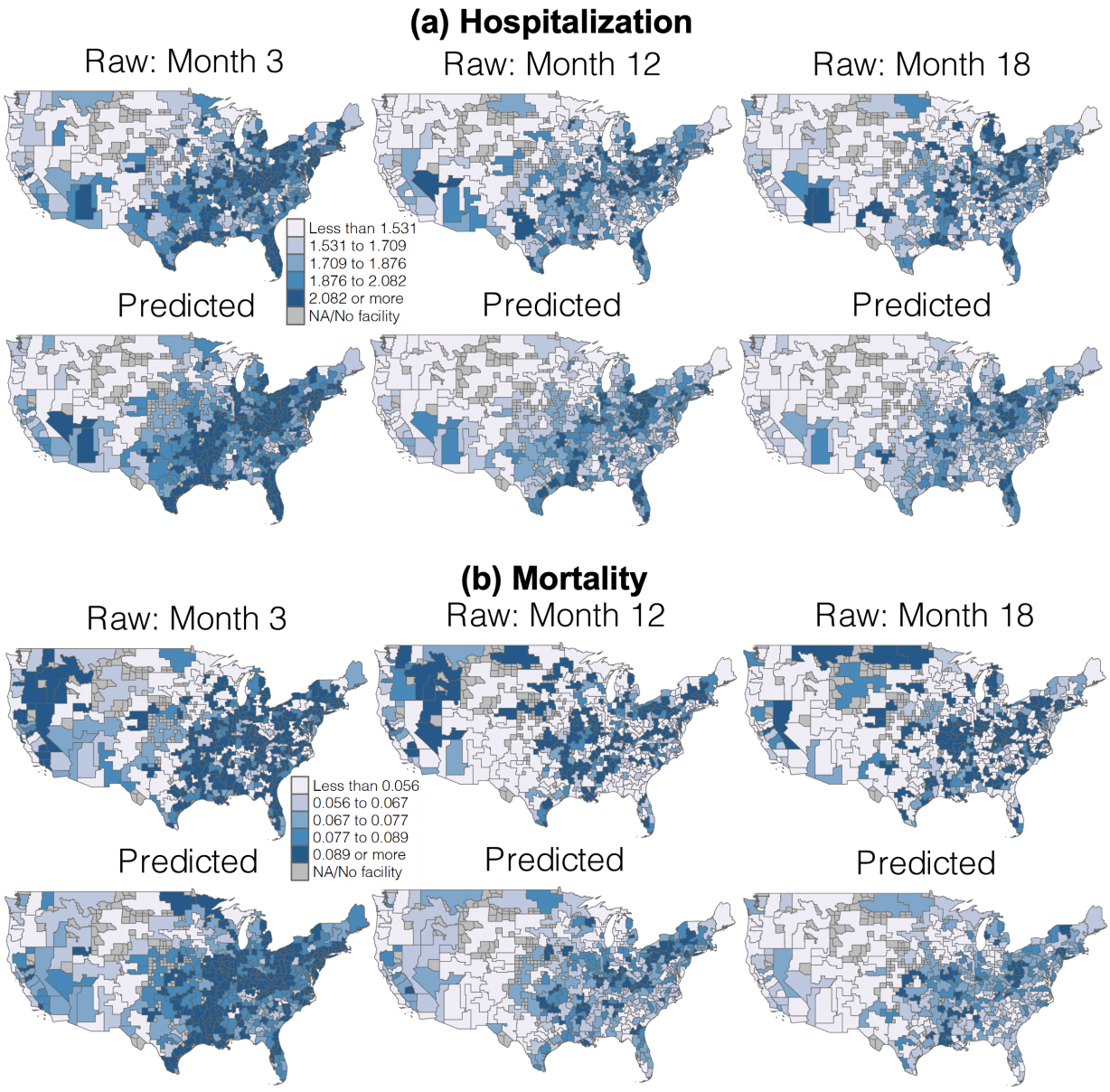

Figure 2.4: The raw and predicted hospitalization (a) and mortality (b) rates from the 3rd, 12th, and 18th months on dialysis for all HSAs.

varying coefficient models that ignore the spatiotemporal correlations in the region-specific deviations using prediction error. The first comparative model is the multivariate varying-coefficient model (MV-VCM) from Step 1 of the MV-VCSTM estimation algorithm, fitted under the working independence assumption. MV-VCM, while accommodating time-varying effects of covariates on the multivariate outcome, ignores the spatiotemporal correlations remaining in the data after adjusting for time-varying effects of covariates. Different from MV-VCM, the second comparative model, referred to as the multivariate varying-coefficient temporal model (MV-VCTM) incorporates temporal trends in the error modeling after adjusting for time-varying effects of covariates: $Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}^{(j)}(t_k) + U_i^{(j)}(t) + \epsilon_i^{(j)}(t_k)$, where $U_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}\psi_\ell^{(j)}(t_k)$ with $\psi_\ell^{(j)}(t)$ denoting the multivariate eigenfunctions and $\rho_{i\ell}$ denoting the corresponding multivariate PC scores. Different from the proposed MV-VCSTM, the PC scores $\rho_{i\ell}$ in MV-VCTM are assumed to be i.i.d across regions, ignoring the spatial correlations in the region-specific deviations. Similar to MV-VCSTM, MV-VCTM is also fitted using Bayesian hierarchical modeling (implemented via WinBUGS), where region-specific deviations are expanded on multivariate eigenfunctions and VCFs are expanded using penalized thin-plate splines. The estimation algorithm for MV-VCTM also starts with an initial MV-VCM fit to the data, but targets the multivariate eigenfunctions by the multivariate functional principal components of Happ and Greven (2018). We compare the three model fits using relative mean squared deviation error (MSDE), i.e., $\text{MSDE}_{\widehat{Y}^{(j)}(t)} = \int\{Y_i^{(j)}(t) - \widehat{Y}_i^{(j)}(t)\}^2\mathrm{d}t / \int\{Y_i^{(j)}(t)\}^2\mathrm{d}t$, where the results are reported in Table 2.1. As expected, MV-VCM, ignoring the spatiotemporal correlation of the error leads to the largest MSDEs, followed by MV-VCTM (which ignores the spatial correlation of the error). By comparison, the proposed MV-VCSTM leads to the smallest MSDEs in modeling both hospitalizations and mortality. As will be shown through the simulation studies summarized in the next section, modeling of the spatiotemporal correlations in the errors via MV-VCSTM, not only leads to smaller prediction error for the outcome trajectories, but it also improves efficiency in inference on the VCFs (evaluated via coverage probability and length of the associated credible bands).

Table 2.1: The mean MSDEs from predicted multivariate response trajectories using MV-VCM, MV-VCTM, MV-VCSTM fits to the USRDS data.

	MV-VCM	**MV-VCTM**	**MV-VCSTM**
Hospitalization	0.048	0.045	0.015
Mortality	0.321	0.314	0.236

2.4 Simulation Studies

Finite sample properties of the proposed estimation and inference algorithm are studied via simulation studies with varying number of regions and measurement error variance for multivariate response from two and three dimensions, and we also explore the performance of the proposed estimation and inference algorithm under aligned and misaligned simulation designs where regions are observed at aligned and misaligned time points, respectively. In addition, we compare the finite sample performance of the proposed MV-VCSTM to that of the two comparative models MV-VCTM and MV-VCM introduced above, which ignore the spatial and spatiotemporal correlations of the error term, respectively.

2.4.1 Simulation Design

In the simulation with the two-dimensional multivariate response, the bivariate response is simulated using

$$Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}^{(j)}(t_k) + \sum_{\ell=1}^{L} \rho_{i\ell}\psi_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k) \quad \text{for} \quad j = 1, 2,$$

where in the aligned design, t_k, $k = 1, \ldots, 24$, denote the equidistant grid of time points between 0 and 1, mimicking the 24 month aligned follow-up in our data application, whereas in the misaligned design, t_k, $k = 1, \ldots, 48$, denote the equidistant grid of time points between 0 and 1, then 50% random missing is induced for each region to achieve region-specific misaligned time points. The covariates $\boldsymbol{X}_i = \{1, X_{i1}, X_{i2}\}^{\mathsf{T}}$, $i = 1, \ldots, n$, include a y-intercept term, where X_{i1}, X_{i2} are generated from a bivariate normal distribution with mean $(0, 0)^{\mathsf{T}}$, variance 1 and covariance $2^{-0.5}$. The time-varying coefficient function $\boldsymbol{\beta}^{(j)}(t_k) =$

$\{\beta_1^{(j)}(t_k), \beta_2^{(j)}(t_k), \beta_3^{(j)}(t_k)\}$, $j = 1, 2$, is set to:

$$\beta_0^{(1)}(t) = \frac{1}{5}\exp(2t-1), \quad \beta_1^{(1)}(t) = \frac{1}{2}t(1-t), \quad \beta_2^{(1)}(t) = \frac{1}{8}(1-t)^2,$$

$$\beta_0^{(2)}(t) = \frac{3}{10}\sqrt{2t} + \frac{1}{10}, \quad \beta_1^{(2)}(t) = \frac{1}{2}(t-0.5)^2, \quad \beta_2^{(2)}(t) = \frac{1}{8}t^2.$$

The multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t_k) = \{\psi_\ell^{(1)}(t_k), \psi_\ell^{(2)}(t_k)\}^\mathsf{T}$, $\ell = 1, 2, 3$ (for $L = 3$), are generated using 11 Fourier basis functions on the interval $\mathcal{T} = [0, 2]$. While $\psi_\ell^{(1)}(t)$, $\ell = 1, 2, 3$, are taken to be the 4th, 5th and 8th Fourier basis functions evaluated on $[0, 1]$, respectively, $\psi_\ell^{(2)}(t)$, $\ell = 1, 2, 3$, are set to equal 4th, 9th and 8th basis functions evaluated on $[1, 2]$, shifted to the left by 1 and multiplied by a random sign. The sign of $\psi_\ell^{(2)}(t)$, $\ell = 1, 2, 3$, are selected to guarantee that the multivariate eigenfunctions are orthonormal on $[0, 1]$. A CAR structure is induced on the multivariate PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\mathsf{T}$. More specifically, the score vector $\boldsymbol{\rho}_\ell$ is generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$, where W denotes the adjacency matrix and D denotes the diagonal matrix with the total number of neighbors for each region, as described in Section 2.2.1. The matrices W and D are specified using the map of $n = 49$ states in the contiguous U.S. (including the District of Columbia) and the map of the $n = 367$ HSAs from our data application. The spatial variance parameters α_ℓ's are taken to be $\alpha_1 = 2.4$, $\alpha_2 = 1.2$ and $\alpha_3 = 0.6$, and the spatial correlation parameter ν is set to 0.97. Finally, the measurement errors $\boldsymbol{\epsilon_i}(t_k) = \{\epsilon_i^{(1)}(t_k), \epsilon_i^{(2)}(t_k)\}^\mathsf{T} \sim N\left((0,0)^\mathsf{T}, \text{diag}(\sigma_1^2, \sigma_2^2)\right)$, where $\sigma_1^2 = \sigma_2^2 = \sigma^2$ equal 0.2 and 2 corresponding to low and high noise simulation cases, respectively. Details on generation of data for the simulation with three-dimensional multivariate response are deferred to Appendix B.2.

2.4.2 Results

The relative mean squared deviation error (MSDE), $\text{MSDE}_{\widehat{\theta}(t)} = [\int\{\widehat{\theta}(t) - \theta(t)\}^2 \text{d}t] / \int\{\theta(t)\}^2 \text{d}t$ (for a generic function $\theta(t)$), and the mean squared error (MSE), $\text{MSE}_{\widehat{\theta}} = (\theta - \widehat{\theta})^2/\theta^2$ (for a generic parameter θ), are utilized to assess estimation of the time-varying and time-invariant

parameters, respectively. In addition to the time-varying model parameters, MSDEs of the estimated region-specific deviations $\widehat{U}_i^{(j)}(t)$ and predicted outcome trajectories $\widehat{Y}_i^{(j)}(t)$ are also reported. Reported summaries exclude outlier MSDE values (MSDE$>$ 5) for $\widehat{U}_i^{(j)}(t)$ with denominator values close to zero (less than 2%). All VCFs are modeled with penalized thin-plate spline basis functions using $Q = 3$ number of knots. Results were found to be robust to varying choice of number of knots in a sensitivity analysis. Details on the specific prior settings used in the Bayesian modeling of Steps 3 and 5 of the estimation algorithm are deferred to Appendix B.3. Reported results are based on a total of 200 Monte Carlo runs. While the MCMC in Step 3 of the estimation algorithm uses two parallel chains with 15000 iterations (5000 burn-in), the MCMC in Step 5 uses one chain with 15000 iterations (5000 burn-in) and returns one out of every ten samples. All Markov chains are verified to have good mixing and convergence properties and the trace plots are given in Appendix B (Figure B.2 and B.3).

The mean MSDE and MSE values from four simulation set-ups with two sets of total number of regions ($n = 49$ and $n = 367$) and two error variances ($\sigma^2 = .2$ and $\sigma^2 = 2.0$) are reported in Tables 2.2 for the two-dimensional multivariate response with aligned time points design. All error measures decrease with increasing number of regions, as expected. In addition, MSDEs for all quantities also decrease with decreasing noise level σ^2, while MSEs for variance components are comparable across the two noise levels, showing that the proposed MV-VCSTM can remove effects of measurement error under both error settings. Even under $n = 49$ and the higher error variance, error measures are small, signaling a good fit. Error measures for $n = 49$ and $\sigma^2 = .2$ and 2.0 from both aligned and misaligned time points designs (reported in Table B.1) and from three-dimensional multivariate outcome with aligned time points design (reported in Table B.2) follow similar trends. Moreover, the estimation of varying coefficient functions $\boldsymbol{\beta}^{(j)}(t)$, multivariate eigenfunctions $\psi_\ell(t)$, as well as variance components (α_ℓ, ν and σ^2) from misaligned time points design is comparable with that from the aligned time points design (Table B.1), indicating the proposed MV-VCSTM methodology is effective with the case where data from different regions are collected at different sets of time points. The estimated VCFs from the simulation runs with the 5th,

50th and 95th percentile MSDEs based on $n = 49$ and $n = 367$ regions are given in Figures B.4 and Figure B.5, respectively. In addition, the estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs based on $n = 49$ and $n = 367$ regions are displayed in Figures B.6 and Figure B.7, respectively. The estimates track the true functions, where quantities are estimated better with increasing number of regions and decreasing error variance, as expected.

Finally, the finite sample performance of the proposed MV-VCSTM is compared to two multivariate varying coefficient models, MV-VCTM and MV-VCM, which ignore the spatial and spatiotemporal correlations of the error term, respectively. Model comparisons at varying error variance for $n = 49$ and $n = 367$ regions are summarized in Tables B.3 and Table 2.3, respectively, for the two-dimensional multivariate response (results for the three-dimensional response from $n = 49$ are deferred to Table B.4). Models are compared with respect to MSDEs of VCFs, region-specific deviations $\widehat{U}_i^{(j)}(t)$, outcome predictions $\widehat{Y}_i^{(j)}(t)$, as well as the VCF coverage probabilities (CP) and the corresponding length of 95% credible bands (CBs). Intervals reported for MV-VCM are approximate 95% pointwise empirical confidence intervals obtained by the `gam()` function in the R package `mgcv`. Simultaneous CBs for MV-VCTM are formed in the same way as MV-VCSTM, based on posterior samples of the VCFs. MV-VCM, ignoring the spatiotemporal correlations in the errors, has the worst performance among the three models, with higher MSDE for the y-intercept VCF and the predicted response trajectories $\widehat{Y}_i^{(j)}(t)$ and significant under coverage for the y-intercept VCF in both dimensions. The proposed MV-VCSTM, incorporating spatiotemporal correlations in the error leads to the lowest MSDEs out of all three methods where length of the CBs are also smaller compared to those from MV-VCTM as well, signaling improved efficiency in inference. Note that under the misaligned time points design (Table B.3), while all three methods (MV-VCM, MV-VCTM, and MV-VCSTM) can deal with misaligned time points design, the proposed MV-VCSTM has the best performance among these three methods, as it does with the aligned time points design. Results given in Table B.4 for the three-dimensional multivariate outcome follow similar trends. Note that CPs reported from MV-VCTM and MV-VCSTM are higher than the nominal level, as expected, as Bayesian CBs tend to be

Table 2.2: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from four simulation settings with two sets of total number of regions and two sets of measurement error variance σ^2 with aligned time points design. Results are based on two-dimensional multivariate response and 200 Monte Carlo runs.

Number of regions, n	**49 regions**		**367 regions**	
Noise level, σ^2	0.2	2	0.2	2
	MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.0033	0.0178	0.0003	0.0015
$\widehat{\beta}_0^{(2)}(t)$	0.0031	0.0236	0.0003	0.0011
$\widehat{\beta}_1^{(1)}(t)$	0.0017	0.0096	0.0004	0.0020
$\widehat{\beta}_1^{(2)}(t)$	0.0013	0.0093	0.0002	0.0015
$\widehat{\beta}_2^{(1)}(t)$	0.0011	0.0126	0.0001	0.0011
$\widehat{\beta}_2^{(2)}(t)$	0.0009	0.0093	0.0001	0.0012
$\widehat{\psi}_1^{(1)}(t)$	0.0795	0.0872	0.0105	0.0179
$\widehat{\psi}_1^{(2)}(t)$	0.0821	0.0846	0.0104	0.0176
$\widehat{\psi}_2^{(1)}(t)$	0.0960	0.1222	0.0164	0.0265
$\widehat{\psi}_2^{(2)}(t)$	0.1060	0.1208	0.0168	0.0281
$\widehat{\psi}_3^{(1)}(t)$	0.0349	0.1697	0.0048	0.0157
$\widehat{\psi}_3^{(2)}(t)$	0.0333	0.1769	0.0050	0.0172
$\widehat{U}_i^{(1)}(t)$	0.0465	0.3447	0.0416	0.2775
$\widehat{U}_i^{(2)}(t)$	0.0470	0.3516	0.0418	0.2789
$\widehat{Y}_i^{(1)}(t)$	0.0251	0.3499	0.0244	0.2010
$\widehat{Y}_i^{(2)}(t)$	0.0205	0.2152	0.0196	0.1594
	MSE			
$\widehat{\alpha}_1$	0.0985	0.1067	0.0284	0.0300
$\widehat{\alpha}_2$	0.0260	0.0347	0.0091	0.0156
$\widehat{\alpha}_3$	0.0188	0.0696	0.0150	0.0246
$\widehat{\nu}$	0.0067	0.0070	0.0066	0.0067
$\widehat{\sigma}^{2(1)}$	0.0022	0.0023	0.0005	0.0005
$\widehat{\sigma}^{2(2)}$	0.0023	0.0024	0.0004	0.0005

conservative (Cox, 1993; Krivobokova et al., 2010).

Table 2.3: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) and length of the 95% credible bands for $n = 367$ regions from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on two-dimensional multivariate response with aligned time points design and 200 Monte Carlo runs.

# regions	**367 regions**					
Noise level	0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
	MSDE					
$\widehat{\beta}_0^{(1)}(t)$	0.0691	0.0513	0.0003	0.0828	0.0606	0.0015
$\widehat{\beta}_0^{(2)}(t)$	0.0690	0.0514	0.0003	0.0829	0.0613	0.0011
$\widehat{\beta}_1^{(1)}(t)$	0.0005	0.0009	0.0004	0.0039	0.0031	0.0020
$\widehat{\beta}_1^{(2)}(t)$	0.0006	0.0008	0.0002	0.0026	0.0026	0.0015
$\widehat{\beta}_2^{(1)}(t)$	0.0012	0.0009	0.0001	0.0017	0.0024	0.0011
$\widehat{\beta}_2^{(2)}(t)$	0.0011	0.0008	0.0001	0.0026	0.0022	0.0012
$\widehat{U}_i^{(1)}(t)$	/	0.2045	0.0416	/	0.4150	0.2775
$\widehat{U}_i^{(2)}(t)$	/	0.2049	0.0418	/	0.4161	0.2789
$\widehat{Y}_i^{(1)}(t)$	0.6748	0.0282	0.0244	0.6777	0.2168	0.2010
$\widehat{Y}_i^{(2)}(t)$	0.5763	0.0222	0.0196	0.5853	0.1734	0.1594
	CP (%)					
$\widehat{\beta}_0^{(1)}(t)$	33.71	100	100	30.06	100	100
$\widehat{\beta}_0^{(2)}(t)$	31.48	100	100	28.08	100	100
$\widehat{\beta}_1^{(1)}(t)$	100	100	100	98.96	97.5	99.5
$\widehat{\beta}_1^{(2)}(t)$	100	100	100	99.98	100	100
$\widehat{\beta}_2^{(1)}(t)$	100	100	100	100	100	100
$\widehat{\beta}_2^{(2)}(t)$	100	100	100	100	100	100
	Length					
$\widehat{\beta}_0^{(1)}(t)$	0.192	1.136	0.955	0.339	1.158	0.920
$\widehat{\beta}_0^{(2)}(t)$	0.196	1.341	0.897	0.349	1.347	0.784
$\widehat{\beta}_1^{(1)}(t)$	0.227	0.412	0.265	0.285	0.557	0.334
$\widehat{\beta}_1^{(2)}(t)$	0.221	0.553	0.428	0.355	0.666	0.427
$\widehat{\beta}_2^{(1)}(t)$	0.213	0.363	0.278	0.277	0.519	0.333
$\widehat{\beta}_2^{(2)}(t)$	0.237	0.472	0.318	0.280	0.601	0.443

2.5 Discussion

A novel multivariate varying coefficient spatiotemporal model (MV-VCSTM) is proposed to study time-varying effects of risk factors jointly on hospitalization and mortality rates in

the U.S. dialysis population. In addition to adjusting for time-varying regression effects of covariates which are characterized as functions of time that patients stay on dialysis, the proposed modeling approach accounts for the remaining joint spatiotemporal patterns of hospitalization and mortality rates among geographic regions (HSAs). To address challenges in estimation and inference in high-dimensional multivariate outcome settings, the proposed estimation relies on multiple dimension reduction techniques. The region-specific random deviations are expanded on multivariate eigenfunctions (which leads to representations in lower dimensions, involving only a few multivariate eigenfunctions), which are targeted using only univariate FPCA expansions, leading to significant computational savings by avoiding smoothing over high dimensional covariance surfaces. Expansions using multivariate eigenfunctions not only help achieve computational savings, they also lead to extra interpretations in modeling, via capturing the leading directions of temporal variation in the errors (after adjustment for time-varying effects of risk factors). Additional computational savings are achieved in MCAR modeling using independent latent Gaussian processes with a CAR correlation structure. Hence, the proposed modeling, while leading to a nonseparable space and time covariance structure in the outcomes, can still easily scale up to multivariate response in higher dimensions. Modeling the spatiotemporal correlations in the errors via MST-VCSTM is shown to improve the efficiency of inference on the VCFs through simulation studies. Finally, applications of the proposed methodology to USRDS data lead to the discovery of significant risk factors (large metropolitan and ADI) for hospitalization rates. In addition, the proposed MV-VCSTM leads to identification of "hot spots" (with high rates in both hospitalization and mortality or with differing patterns in the outcomes) and time periods with elevated rates after transitioning to dialysis.

CHAPTER 3

Bayesian Multivariate Joint Modeling of Longitudinal, Recurrent, and Competing Risk Terminal Events in Patients with Chronic Kidney Disease

3.1 Introduction

Chronic kidney disease (CKD), defined by having an estimated glomerular filtration rate (eGFR) less than 60 ml/min/1.73m^2 or a urine albumin-to-creatinine ratio of 30 mg/g or higher, affects more than 1 in 7 U.S. adults (USRDS, 2022). In addition, as of 2020, end-stage kidney disease (ESKD), the final stage of CKD, impacted almost 808,000 individuals in the U.S., with patients typically dependent on the life-sustaining treatment of dialysis for the duration of their lives or until kidney transplantation. Recent studies have highlighted a significant association between reduced eGFR and increased risk of mortality, cardiovascular (CV) events, and hospitalization among CKD patients (Go et al., 2004). Moreover, individuals with ESKD undergoing hemodialysis exhibit an even higher prevalence of cardiovascular disease (CVD), at about 77% (USRDS, 2022). As a leading cause of mortality, morbidity and frequent hospitalizations, CKD poses a considerable economic burden which escalates substantially with increasing disease severity (Go et al., 2004). Hence, there exists an urgent imperative to develop an effective monitoring tool for CKD progression, along with modeling of interdependent outcomes, recurrent CV hospitalizations and multiple terminal events (ESKD and death), and associated risk factors.

To gain deeper insights into the risk factors that influence the progression of CKD and CVD, as well as the associated morbidity and mortality in CKD patients, the National

Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) established the Chronic Renal Insufficiency Cohort (CRIC) study in 2001 (Feldman et al., 2003). This ongoing prospective cohort study recruits adult patients who are aged 21 to 74 years old with mild to moderate CKD. Inspired by the broad aims of the CRIC study, we introduce a novel Bayesian multivariate joint model (BM-JM), designed to study the interdependent dynamics of CKD progression defined by longitudinal eGFR trajectories, recurrent hospitalizations due to CV events, and competing-risk terminal outcomes of ESKD and death. The proposed BM-JM also facilitates dynamic prediction of survival probabilities of each competing-risk terminal outcome for future subjects based on comprehensive individual data, including baseline characteristics, a complete history of previous longitudinal eGFR measurements and past recurrent CV events. Note that even though ESKD or kidney failure may not strictly qualify as a "terminal" event in clinical terms, it is commonly modeled as a survival outcome in the nephrology literature (Yang et al., 2014). Since the longitudinal follow-up of eGFR ends upon the occurrence of ESKD (defined as the initiation of dialysis or kidney transplant) in the CRIC study, it is also treated as a competing risk in our BM-JM modeling.

Methodological development and clinical applications for joint modeling have grown substantially over the past two decades, as extensively reviewed in Gould et al. (2015). A commonly used approach to induce dependency among the multiple outcomes modeled is the inclusion of shared frailty/random effects terms (Wulfsohn and Tsiatis, 1997). While most of the literature is on modeling of a continuous longitudinal outcome and a right-censored event time, the joint modeling has been extended to handle interval- (Gueorguieva et al., 2012) and left-censored (Thiébaut et al., 2005) data, discrete event times (Albert and Shih, 2010), competing risks (Williamson et al., 2008), and recurrent event times (Ren et al., 2021). The extension of joint models to accommodate multiple event time outcomes is particularly valuable in health research and clinical decision-making, as patients may often experience multiple, recurrent, or a succession of clinical events (e.g. recurrent CV-related hospitalizations in the CKD population).

Another important extension is the use of joint modeling for dynamic prediction of survival probabilities, based on all available baseline characteristics, longitudinal and/or recur-

rent events data. Dynamic predictions offer personalized forecasts that can be continuously updated as additional longitudinal and/or recurrent events data become available. Instead of the conventional dependency formulation through random effects, recent literature on dynamic prediction methods include direct longitudinal biomarker effects in the survival submodel (Rizopoulos, 2011). This approach allows for the direct evaluation of the predictive value of longitudinal biomarker measurements, while also enhancing the accuracy of the dynamic predictions.

Despite important extensions in recent years, existing literature on joint modeling of longitudinal measurements and multiple event time outcomes that include both recurrent and terminal events, remains limited. The literature includes a joint model of (i) CD4 cell counts, intensity of opportunistic disease, and death in patients with HIV (Liu and Huang, 2009); (ii) tumor size, occurrence of new lesions, and death in patients with metastatic colorectal cancer (Król et al., 2016); and (iii) eGFR trajectories, recurrent CV events and a composite terminal event (ESKD or death) in CKD patients (Kürüm et al., 2024). All three models utilize shared random effects to account for the interdependency among the trivariate outcomes, and are limited to a single terminal event outcome.

To address the existing modeling gaps in multivariate joint modeling of longitudinal trajectories, recurrent events and competing-risk terminal events as well as developing an efficient dynamic prediction procedure for competing-risk terminal outcomes utilizing the combined history of longitudinal measurements and recurrent events, we propose a novel Bayesian multivariate joint model (BM-JM). By explicitly modeling the effects of longitudinal biomarker in the competing-risk survival submodels, the proposed BM-JM allows for direct quantification and interpretation of the predictive value of the longitudinal measurements on competing-risk terminal outcomes, while also enabling more precise and comprehensive dynamic predictions for future patients. The proposed framework allows for prediction of each competing risk individually, utilizing the most recent available history of not only longitudinal measurements but also recurrent events. Note that while our data application illustrates the proposed model for longitudinal, recurrent, and two competing-risk terminal processes, the BM-JM framework can be extended to accommodate more complex

interdependent multi-disease applications.

The remainder of the paper is organized as follows. Details of the proposed BM-JM, including the proposed estimation and inference procedures, are outlined in Section 3.2. Section 3.3 introduces the proposed individualized dynamic prediction framework as well as metrics for evaluating predictive accuracy. Application of the proposed methodology to the CRIC study for joint modeling of kidney function decline, recurrent CV events, and competing-risk terminal events (ESKD and death) for patients with CKD is given in Section 3.4. Comprehensive simulations to study the finite sample properties of the proposed BM-JM, including its predictive performance are outlined in Section 3.5. In both the CRIC data application and the simulation studies, the proposed BM-JM is compared with three marginal models (i.e., the longitudinal model (L-M), the recurrent model (R-M), and the competing-risk terminal model (T-M)) and two simplified joint models (i.e., the joint model of longitudinal measurements and competing-risk terminal events (LT-JM) and the joint model of recurrent and competing-risk terminal events (RT-JM)). These comparative analyses reveal that ignoring either recurrent events or longitudinal measurements in joint modeling adversely impacts both parameter inference and predictive performance. We conclude with a brief discussion of key findings and potential future works in Section 3.6.

3.2 Proposed Bayesian Multivariate Joint Model (BM-JM)

3.2.1 Model Specification

Chronic kidney disease is commonly defined based on eGFR, which serves as a widely accepted measure of kidney function (Go et al., 2004). Therefore, eGFR measurements are chosen as the longitudinal biomarker in our data application. Let $Y_i(t)$ denote the value of the longitudinal biomarker (i.e., eGFR trajectory) for the ith subject ($i = 1, \ldots, n$) at time t. The vector of longitudinal measurements, $\mathbf{Y}_i = \{Y_i(t_{i1}), \ldots, Y_i(t_{iK_i})\}^\mathsf{T}$, is observed at subject-specific time points t_{ik}, $k = 1, \ldots, K_i$. The true but unobserved values of the longitudinal biomarker over time are denoted by $\xi_i(t)$. A linear structure for $\xi_i(t)$ was selected as

it is a common practice in kidney disease modeling. Longitudinal studies have consistently demonstrated that eGFR trajectories typically decline linearly in CKD patients, allowing linear models to effectively estimate both the rate of decline and its clinical significance (Coresh et al., 2007; Skupien et al., 2016). Additionally, adopting a linear structure for $\xi_i(t)$ ensures simplicity and enhances interpretability within the joint modeling framework.

Patients with reduced kidney function face significantly higher risks of mortality and kidney failure compared to the general population (Matsushita et al., 2010). Consequently, ESKD and mortality are commonly used as time-to-event outcomes in kidney disease studies (Yang et al., 2014; Fine and Gray, 1999; Grams et al., 2013). In the CRIC study, longitudinal follow-up of eGFR terminates at either ESKD or death, making these two outcomes competing-risk terminal events. Let $T_i^{*(w)}$ $(w = 1, \ldots, W)$ denote the true failure time for the wth competing terminal event for subject i, and C_i the censoring time. For simplicity, $W = 2$ is used in our developments, corresponding to ESKD and death, although the proposed framework can be extended to accommodate multiple competing risk terminal events $(W \geq 3)$. The observed terminal event time, T_i, is defined as the minimum of the true failure times and the potential censoring time C_i, i.e., $T_i = \min(T_i^{*(1)}, T_i^{*(2)}, C_i)$, with the event indicator δ_i taking values of 0, 1, or 2 to denote censoring, ESKD and death, respectively.

CKD is also a recognized risk factor for recurrent CV events (Matsushita et al., 2010), which can exacerbate renal insufficiency and increase mortality risk for CKD patients (Muntner et al., 2002). For the recurrent event processes (e.g., recurrent CV events in our data application), let G_{ij}^* denote the jth true gap time, defined as the time between recurrent events $j-1$ and j. Further let $R_{ij}^* = G_{i1}^* + \ldots + G_{ij}^*$ denote the true total time from the study onset to event j, and let $C_{ij} = C_i - R_{i,j-1}^*$ denote the jth gap censoring time for $j = 1, \ldots, J_i$, with J_i denoting the total number of recurrent events of subject i before censoring. The observed gap time and recurrent event indicator are then defined as $G_{ij} = \min(G_{ij}^*, C_{ij})$ and $\lambda_{ij} = I_{(G_{ij} \leq C_{ij})}$, respectively, with $I_{(\cdot)}$ denoting the indicator function.

Several studies have employed joint models to explore the interdependency among longitudinal eGFR measurements, recurrent CV events, and terminal outcomes, such as ESKD and death, in CKD patients. For example, Hillege et al. (2006) utilized joint models to study

the relationship between eGFR decline and CV event outcomes in this population. More recently, Kürüm et al. (2024) introduced a Bayesian trivariate model that integrates longitudinal kidney disease progression, recurrent CV events, and a composite terminal event to examine their interrelationships. Despite these advances, existing models exhibit certain limitations. They often do not fully account for competing-risk terminal processes or provide tools for dynamically predicting terminal outcomes using the combined history of longitudinal eGFR trajectories and recurrent CV events. To address these gaps, we propose the BM-JM, a Bayesian joint model that integrates longitudinal eGFR measurements, recurrent CV events, and competing-risk terminal events. The BM-JM explicitly captures the interrelationships among these processes and facilitates dynamic prediction of survival probabilities for both ESKD and death in new patients by incorporating their updated longitudinal and recurrent event profiles. The specific formulation of the proposed BM-JM is detailed below:

$$
\begin{aligned}
Y_i(t) &= \xi_i(t) + \varepsilon_i(t) = \mathbf{X}_i^{\mathsf{T}}\boldsymbol{\beta}_\ell + \mathbf{Z}_i^{\mathsf{T}}\boldsymbol{\phi}_\ell + \gamma t + b_{i0} + b_{i1}t + \varepsilon_i(t), \\
r_{ij}(g_{ij}) &= h_{r0}(g_{ij}) \exp\left(\mathbf{X}_i^{\mathsf{T}}\boldsymbol{\beta}_r + \mathbf{Z}_i^{\mathsf{T}}\boldsymbol{\phi}_{rj} + \sum_{m=0}^{j-1}\alpha_m + \eta_{r0}b_{i0} + \eta_{r1}b_{i1} + \nu_i\right), \qquad (3.1) \\
h_i^{(1)}(t) &= h_{t0}^{(1)}(t) \exp\left(\mathbf{X}_i^{\mathsf{T}}\boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^{\mathsf{T}}\boldsymbol{\phi}_t^{(1)} + \eta_t^{(1)}\xi_i(t) + \zeta^{(1)}\nu_i\right), \\
h_i^{(2)}(t) &= h_{t0}^{(2)}(t) \exp\left(\mathbf{X}_i^{\mathsf{T}}\boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^{\mathsf{T}}\boldsymbol{\phi}_t^{(2)} + \eta_t^{(2)}\xi_i(t) + \zeta^{(2)}\nu_i\right),
\end{aligned}
$$

where $Y_i(t)$, $r_{ij}(g_{ij})$ and $h_i^{(w)}(t)$ represent the observed longitudinal eGFR measurements and the hazard functions for the j-th recurrent and w-th competing-risk terminal events for the i-th subject, respectively. Note that different time indices are utilized across the submodels in (3.1). The longitudinal and competing-risk terminal submodels are parameterized using follow-up time indexed by t, whereas the recurrent event submodel employs gap time, indexed by g_{ij}, which refers to the time interval between the $(j-1)$-th and j-th recurrent events for the i-th subject.

For clarity, we separate the covariates into two groups: primary covariates (primary risk factors) and secondary covariates (secondary risk factors). The primary covariates, denoted by $\mathbf{Z}_i = (Z_{i1}, \ldots, Z_{id})^{\mathsf{T}}$, are the main risk factors of interest. In our data application, these

include diabetes, hypertension and history of CVD, which are key contributors to morbidity and mortality in CKD patients. The recurrent event submodel captures how these primary risk factors influence the occurrence of the first, second, and subsequent CV events, with these effects modeled through event-varying coefficients $\boldsymbol{\phi}_{rj}$. The secondary risk factors, denoted by $\mathbf{X}_i = (X_{i1}, \ldots, X_{id'})^\mathsf{T}$, comprise a broader set of covariates, including demographic variables, comorbidities and laboratory measurements. These covariates account for additional variability and allow for a more comprehensive characterization of the patient population. The regression coefficients for these secondary and primary covariates are denoted by $\boldsymbol{\beta} = (\boldsymbol{\beta}_\ell^\mathsf{T}, \boldsymbol{\beta}_r^\mathsf{T}, \boldsymbol{\beta}_t^{(1)\mathsf{T}}, \boldsymbol{\beta}_t^{(2)\mathsf{T}})$ and $\boldsymbol{\phi} = (\boldsymbol{\phi}_\ell^\mathsf{T}, \boldsymbol{\phi}_r^\mathsf{T}, \boldsymbol{\phi}_t^{(1)\mathsf{T}}, \boldsymbol{\phi}_t^{(2)\mathsf{T}})$, respectively, where $\boldsymbol{\phi}_r = (\boldsymbol{\phi}_{r1}^\mathsf{T}, \ldots, \boldsymbol{\phi}_{rJ_{\max}}^\mathsf{T})^\mathsf{T}$ denotes the event-varying coefficient vector of the recurrent event submodel, and $J_{\max}$ denotes the maximum number of recurrent events among all subjects, i.e., $J_{\max} = \max(J_1, \ldots, J_n)$. Note that although all the submodels in (3.1) include a common set of covariates, the proposed joint model can easily be extended to accommodate design vectors with different dimensionality and composition for each submodel separately.

In the proposed recurrent event submodel, the recurrent event effect is modeled not only via $\boldsymbol{\phi}_{rj}$, the event-varying effect of covariates, but also via α_m, the additional risk introduced by the mth event (with $m = 1, \ldots, J_{\max} - 1$, and $\alpha_0 = 0$), where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{J_{\max}-1})$ denotes the event effect vector. Hence, $\exp(\alpha_m)$ is the hazard ratio of having an event (e.g., next CV event) for a patient with m previous events relative to a patient with $(m-1)$ previous events.

The interdependency among the submodels within the proposed joint model is established through shared random effects and the true underlying values of longitudinal biomarker, $\xi_i(t)$. Specifically, the dependency between the longitudinal and recurrent outcomes is captured via random effects: the random intercept b_{i0} and random slope b_{i1}, denoted as $\mathbf{b}_i = (b_{i0}, b_{i1})^\mathsf{T}$, and their association is quantified by the regression parameters η_{r0} and η_{r1}. Additionally, a frailty term, ν_i, is incorporated into the model to account for unobserved subject-specific heterogeneity that is not explained by the covariates or the longitudinal process $\xi_i(t)$, as well as the dependency among recurrent events. This frailty term also links the recurrent outcomes to competing risks, with the strength of this association described by the parameter

$\zeta^{(w)}$. Furthermore, each competing risk w is assumed to depend on the true but unobserved values of the longitudinal biomarker, $\xi_i(t)$, which serves as a crucial link between the longitudinal and the w-th competing-risk terminal processes. This dependency is captured through the regression parameter $\eta_t^{(w)}$, which enables direct interpretation of the biomarker's predictive impact on each competing risk and allows comparisons of its effects across different competing risks. In addition, in the proposed BM-JM, the entire longitudinal measurement profile up to the prediction time t is integrated into the competing-risk terminal submodels, alongside the frailty term ν_i from the recurrent event submodel. This framework facilitates flexible and dynamic survival probability predictions for competing-risk terminal outcomes. Notably, ν_i and $\xi_i(t)$ serve distinct purposes in the joint modeling. The frailty term ν_i captures unobserved subject-specific heterogeneity and the dependency among recurrent events. Additionally, ν_i acts as a bridge linking the recurrent events to competing-risk hazards. In contrast, $\xi_i(t)$ represents the true underlying values of the longitudinal biomarker and directly associate the longitudinal outcomes with the competing-risk hazards. By modeling ν_i and $\xi_i(t)$ separately, the proposed framework achieves clearer interpretability of their respective roles and contributions to the joint model.

The vector of combined random effects is denoted by $\mathbf{u}_i = (\mathbf{b}_i^\mathsf{T}, \nu_i)^\mathsf{T}$ where $\mathbf{b}_i^\mathsf{T}$ and ν_i are assumed to be independent, and $\mathbf{u}_i$ follows a multivariate normal distribution $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = [\mathbf{\Sigma}_b \ \mathbf{0}; \mathbf{0} \ \sigma_\nu^2]$, $\mathbf{\Sigma}_b = \left[\sigma_{b_0}^2 \ \sigma_{01}; \sigma_{01} \ \sigma_{b_1}^2\right]$, $\sigma_{01} = \rho_b \sigma_{b_0} \sigma_{b_1}$, and ρ_b denoting the correlation between the random intercept and slope terms. In addition, $h_{r0}(g_{ij})$ denotes the baseline hazard function for the j-th recurrent process of the i-th subject, while $h_{t0}^{(w)}(t)$ represents the baseline hazard function for the wth competing-risk terminal process. The error term $\varepsilon_i(t)$ is assumed to be independent across time and follows a normal distribution $\varepsilon_i(t) \sim N(0, \sigma_\varepsilon^2)$. Furthermore, the error term is assumed to be mutually independent of the random effects. Finally, note that in our framework, longitudinal measurements and recurrent events can be observed at the same time; however, neither can be observed after the observed terminal event time T_i. This is consistent with the CRIC study, where eGFR measurements end at ESKD (not meaningful after ESKD) and obviously at death. Furthermore, a noninformative censoring mechanism is assumed for the terminal event, which ensures that the censoring

mechanism for terminal events does not carry predictive information regarding the longitudinal process or the covariates.

3.2.2 Bayesian Estimation and Inference

To estimate the parameters of our joint model, we propose a Bayesian formulation and derive posterior inferences using a Markov Chain Monte Carlo (MCMC) algorithm. Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\eta}, \boldsymbol{\zeta}, \boldsymbol{\sigma}^2, \rho_b, \boldsymbol{\psi}_{h_0})^\mathsf{T}$ denote the vector of parameters from all submodels with $\boldsymbol{\beta} = (\boldsymbol{\beta}_\ell^\mathsf{T}, \boldsymbol{\beta}_r^\mathsf{T}, \boldsymbol{\beta}_t^{(1)\mathsf{T}}, \boldsymbol{\beta}_t^{(2)\mathsf{T}})$, $\boldsymbol{\phi} = (\boldsymbol{\phi}_\ell^\mathsf{T}, \boldsymbol{\phi}_r^\mathsf{T}, \boldsymbol{\phi}_t^{(1)\mathsf{T}}, \boldsymbol{\phi}_t^{(2)\mathsf{T}})$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{J_{max}-1})$, $\boldsymbol{\eta} = (\eta_{r0}, \eta_{r1}, \eta_t^{(1)}, \eta_t^{(2)})$, $\boldsymbol{\zeta} = (\zeta^{(1)}, \zeta^{(2)})$, $\boldsymbol{\sigma}^2 = (\sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_\varepsilon^2, \sigma_\nu^2)$, and $\boldsymbol{\psi}_{h_0}$ denoting the coefficients used to model the baseline hazard functions. The joint likelihood of the model is derived under the conditional independence assumption, that is, the random effects (REs: $\mathbf{u}_i = (\mathbf{b}_i^\mathsf{T}, \nu_i)^\mathsf{T}$) account for all dependencies among the multivariate outcomes, and that given the REs, the longitudinal, recurrent and competing-risk terminal processes are assumed to be independent. Conditional independence assumption leads to the conditional likelihood

$$\begin{aligned} p(\mathbf{Y}_i, \mathbf{G}_i, \boldsymbol{\lambda}_i, T_i, \delta_i \mid \mathbf{u}_i; \boldsymbol{\theta}) &= p(\mathbf{Y}_i \mid \mathbf{b}_i; \boldsymbol{\theta})\; p(\mathbf{G}_i, \boldsymbol{\lambda}_i \mid \mathbf{u}_i; \boldsymbol{\theta})\; p(T_i, \delta_i \mid \xi_i(\cdot), \nu_i; \boldsymbol{\theta}) \\ &= \prod_{k=1}^{K_i} p(Y_{ik} \mid \mathbf{b}_i; \boldsymbol{\theta}) \prod_{j=1}^{J_i} p(G_{ij}, \lambda_{ij} \mid \mathbf{u}_i; \boldsymbol{\theta})\; p(T_i, \delta_i \mid \xi_i(\cdot), \nu_i; \boldsymbol{\theta}), \end{aligned}$$

where for the ith patient, $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{iK_i})^\mathsf{T}$ denotes the $K_i \times 1$ vector of longitudinal outcomes with $Y_{ik} = Y_i(t_{ik})$, $k = 1, \ldots, K_i$, and $\mathbf{G}_i = (G_{i1}, \ldots, G_{iJ_i})^\mathsf{T}$ and $\boldsymbol{\lambda}_i = (\lambda_{i1}, \ldots, \lambda_{iJ_i})^\mathsf{T}$ are the $J_i \times 1$ vectors of recurrent event times and recurrent event indicators, respectively. Hence, the posterior distribution is given by

$$\begin{aligned} p(\boldsymbol{\theta}, \mathbf{u}_i \mid \mathbf{Y}_i, \mathbf{G}_i, \boldsymbol{\lambda}_i, T_i, \delta_i) &\propto p(\mathbf{Y}_i, \mathbf{G}_i, \boldsymbol{\lambda}_i, T_i, \delta_i \mid \mathbf{u}_i, \boldsymbol{\theta})\; p(\mathbf{u}_i, \boldsymbol{\theta}) \\ &\propto p(\mathbf{Y}_i \mid \mathbf{b}_i; \boldsymbol{\theta})\; p(\mathbf{G}_i, \boldsymbol{\lambda}_i \mid \mathbf{u}_i; \boldsymbol{\theta})\; p(T_i, \delta_i \mid \xi_i(\cdot), \nu_i; \boldsymbol{\theta})\; p(\mathbf{u}_i \mid \boldsymbol{\theta})\; p(\boldsymbol{\theta}). \end{aligned}$$

In our data application, where the longitudinal outcome is continuous, the likelihood contributions from each submodel is given by

$$p(\mathbf{Y}_i \mid \mathbf{b}_i, \boldsymbol{\theta}) = \prod_{k=1}^{K_i} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[\frac{\{Y_{ik} - (\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_\ell + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_\ell + \gamma t_{ik} + b_{i0} + b_{i1}t_{ik})\}^2}{2\sigma_\varepsilon^2}\right],$$

$$p(\mathbf{G}_i, \boldsymbol{\lambda}_i \mid \mathbf{u}_i, \boldsymbol{\theta}) = \prod_{j=1}^{J_i} \left[\left\{h_{r0}(G_{ij}) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_r + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_{rj} + \sum_{m=0}^{j-1}\alpha_m + \eta_{r0}b_{i0} + \eta_{r1}b_{i1} + \nu_i\right)\right\}^{\lambda_{ij}}\right.$$
$$\left.\times \exp\left\{-\int_0^{G_{ij}} h_{r0}(g) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_r + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_{rj} + \sum_{m=0}^{j-1}\alpha_m + \eta_{r0}b_{i0} + \eta_{r1}b_{i1} + \nu_i\right) \mathrm{d}g\right\}\right], \quad (3.2)$$

and

$$p(T_i, \delta_i \mid \xi_i(\cdot), \nu_i, \boldsymbol{\theta}) = \prod_{w=1}^{2} \left[\left\{h_{t0}^{(w)}(T_i) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(w)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(w)} + \eta_t^{(w)}\xi_i(T_i) + \zeta^{(w)}\nu_i\right)\right\}^{I(\delta_i = w)}\right.$$
$$\left.\times \exp\left\{-\int_0^{T_i} h_{t0}^{(w)}(v) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(w)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(w)} + \eta_t^{(w)}\xi_i(v) + \zeta^{(w)}\nu_i\right) \mathrm{d}v\right\}\right], \quad (3.3)$$

where $\xi_i(t) = \mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_\ell + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_\ell + \gamma t + b_{i0} + b_{i1}t$ denotes the value of the true but unobserved biomarker for subject i at time point t. Note that the integrals in (3.2) and (3.3) do not have closed-form solutions and the 15-point Gauss-Kronrod quadrature rule (Kahaner et al., 1989) is employed to approximate the integrals.

Bayesian P-splines with a relatively large number of equally spaced knots are employed in estimation of the baseline hazard functions. To avoid overfitting, a roughness penalty is incorporated as proposed by Eilers and Marx (1996). Specifically, let $h_0(t)$ denote any baseline hazard function in the recurrent or competing-risk event submodels. It is targeted via $\log\{h_0(t)\} = \sum_{q=1}^{Q} \psi_{h_0,q} B_q(t)$, where $B_q(t)$ denotes the B-spline basis functions and $\boldsymbol{\psi}_{h_0} = \{\psi_{h_0,1}, \ldots, \psi_{h_0,Q}\}^\mathsf{T}$ denotes the corresponding Q-dimensional vector of spline coefficients. Note that while we use Bayesian P-splines with equal number of knots in the estimation of all baseline hazard functions in our data application, our proposed estimation and inference procedure can accommodate both parametric and other nonparametric forms for baseline hazard functions, and allow for the adoption of varying numbers of knots within each baseline hazard function across different submodels.

The priors for fixed-effects parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_\ell^\mathsf{T}, \boldsymbol{\beta}_r^\mathsf{T}, \boldsymbol{\beta}_t^{(1)\mathsf{T}}, \boldsymbol{\beta}_t^{(2)\mathsf{T}})$, $\boldsymbol{\phi} = (\boldsymbol{\phi}_\ell^\mathsf{T}, \boldsymbol{\phi}_r^\mathsf{T}, \boldsymbol{\phi}_t^{(1)\mathsf{T}}, \boldsymbol{\phi}_t^{(2)\mathsf{T}})$, γ, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{J_{\max}-1})$, and the association parameters $\boldsymbol{\eta} = (\eta_{r0}, \eta_{r1}, \eta_t^{(1)}, \eta_t^{(2)})$ and $\boldsymbol{\zeta} = (\zeta^{(1)}, \zeta^{(2)})$, are taken to be normal with mean zero and variance $\sigma_\star^2$, where the hyperparameters are set as inverse gamma (IG): $\sigma_\star^2 \sim \text{IG}(a_{1\star}, a_{2\star})$. IG priors are also assumed for the variance terms $\boldsymbol{\sigma}^2 = (\sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_\varepsilon^2, \sigma_\nu^2)$, i.e., $\sigma_*^2 \sim \text{IG}(a_{1*}, a_{2*})$ ($*$ denoting b_0, b_1, ε or ν), and a uniform prior is assumed for ρ_b, the correlation between the random intercept and slope terms $\mathbf{b}_i = (b_{i0}, b_{i1})^\mathsf{T}$. Note that we assign the same IG prior for the variance terms in $\boldsymbol{\sigma}^2$, however, our method can accommodate different priors for each variance term if needed. In the estimation of the baseline hazard functions, normal priors are used for the spline coefficients, such that $\boldsymbol{\psi}_{h_0}|\kappa_{h_0} \sim N(0, \kappa_{h_0}\mathcal{P}_{h_0})$, where κ_{h_0} denotes the variance parameter controlling the smoothness of the baseline hazard functions, and $\mathcal{P}_{h_0}$ denotes the penalty matrix. Similar to the variance terms, IG priors are used for the variance parameter κ_{h_0}, $\kappa_{h_0} \sim \text{IG}(a_{1\kappa}, a_{2\kappa})$, where $a_{1\kappa}$ is set to 1 and $a_{2\kappa}$ to a small number, following Jullion and Lambert (2007). The penalty matrix $\mathcal{P}_{h_0}$ is computed using the second-order difference matrix Ω of dimension $Q \times Q$ in our applications, i.e., $\mathcal{P}_{h_0} = \Omega^\mathsf{T}\Omega + 10^{-4}I$, with I representing the $Q \times Q$ identity matrix. The addition of a small positive constant (10^{-4}) to the diagonal elements of the product $\Omega^\mathsf{T}\Omega$ helps stabilize the computation and prevents issues related to singularity or numerical instability (Eilers and Marx, 1996). Specific values for these priors are detailed in Appendix C.1. Finally, pointwise credible intervals (CIs) are utilized for inference based on the posterior sample. Let τ denote a parameter within the full parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\eta}, \boldsymbol{\zeta}, \boldsymbol{\sigma}, \rho_b, \boldsymbol{\psi}_{h_0})^\mathsf{T}$, and $\widehat{\tau}$ and $\text{SD}(\tau)$ denote the mean and standard deviation of τ obtained based on the MCMC samples, respectively. Then the $(1-\alpha)$ pointwise credible intervals are given by $\left[\widehat{\tau} \pm \Phi_{\alpha/2}\text{SD}(\tau)\right]$, where $\Phi_{\alpha/2}$ denotes the $100 \times (1-\alpha/2)$-percentile of the standard normal distribution.

3.3 Dynamic Prediction of Cumulative Incidence Functions

3.3.1 Proposed Prediction Algorithm

For a BM-JM fit using a training dataset $\mathcal{D}_n = \{\mathbf{Y}_i, \mathbf{G}_i, \boldsymbol{\lambda}_i, T_i, \delta_i; i = 1, \ldots, n\}$ with n subjects, we consider predicting survival probabilities for each competing-risk terminal outcome on an independent testing dataset of n_p subjects. More specifically, the cumulative incidence probability of each competing-risk terminal outcome is targeted for a new patient p ($p = 1, \ldots, n_p$) with baseline characteristics, a set of longitudinal measurements $\mathcal{Y}_p(t) = \{Y_p(t_{pk}) : 0 < t_{pk} \leq t,\ k = 1, \ldots, K\}$ and recurrent events history $\mathcal{R}_p(t) = \{\{G_{pj}, \lambda_{pj}\} : 0 < \sum_{j=1}^{J} G_{pj} \leq t, j = 1, \ldots, J\}$ till the prediction time t. Given that patient p has not experienced a competing-risk terminal event up to time t, the goal is to predict the subject-p-specific cumulative incidence probability for patient p to experience the wth competing-risk terminal event within the time interval $(t, s]$ given all information available till time t,

$$\pi_p^{(w)}(t, s) \ = \ P\left(t < T_p^{*(w)} < s \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t), \mathcal{D}_n\right),$$

where $s = t + \Delta t$ with $\Delta t > 0$ denoting the prediction window. In this definition, we assume that the previous longitudinal measurements, as well as recurrent events are directly related to the competing-risk failure mechanism.

Note that these predictions are dynamic in the sense that they change with increasing prediction time t and available individual longitudinal and recurrent events information. Under the Bayesian formulation of the joint model, the estimation of $\pi_p^{(w)}(t, s)$ is based on the corresponding posterior predictive distributions, namely,

$$\pi_p^{(w)}(t, s) \ = \ \int P\left(t < T_p^{*(w)} < s \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \times p(\boldsymbol{\theta} \mid \mathcal{D}_n) \mathrm{d}\boldsymbol{\theta}. \quad (3.4)$$

The second term of the integrand in (3.4), $p(\boldsymbol{\theta} \mid \mathcal{D}_n)$, is the posterior distribution of the parameters given the observed data $\mathcal{D}_n$. The first term of the integrand in (3.4) can be given

as

$$\int P\left(t < T_p^{*(w)} < s \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t), \mathbf{u}_p; \boldsymbol{\theta}\right) \times p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \mathrm{d}\mathbf{u}_p. \tag{3.5}$$

Using the full conditional independence assumptions, that is, the random effects $\mathbf{u}_p$ account for all dependencies among the multivariate outcomes, (3.5) can further be simplified as

$$\int P\left(t < T_p^{*(w)} < s \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathbf{u}_p; \boldsymbol{\theta}\right) \times p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \mathrm{d}\mathbf{u}_p.$$

Furthermore, by the definition of conditional probability, (3.5) can be expressed as

$$\begin{aligned} & \int \frac{P\left(t < T_p^{*(w)} < s, \cup_{w=1}^{2}(T_p^{*(w)} > t) \mid \mathbf{u}_p; \boldsymbol{\theta}\right)}{P\left(\cup_{w=1}^{2}(T_p^{*(w)} > t) \mid \mathbf{u}_p; \boldsymbol{\theta}\right)} \; p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \mathrm{d}\mathbf{u}_p \\ \equiv & \int \frac{\mathrm{CIF}_p^{(w)}(t,s)}{\mathrm{S}_p(t)} \times p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \mathrm{d}\mathbf{u}_p, \end{aligned} \tag{3.6}$$

where

$$\mathrm{S}_p(t) = \prod_{w=1}^{2}\left[\exp\left\{-\int_0^t h_{t0}^{(w)}(v)\exp\left(\mathbf{X}_p^{\intercal}\boldsymbol{\beta}_t^{(w)} + \mathbf{Z}_p^{\intercal}\boldsymbol{\phi}_t^{(w)} + \eta_t^{(w)}\xi_p(v) + \zeta^{(w)}\nu_p\right)\mathrm{d}v\right\}\right]$$

and

$$\mathrm{CIF}_p^{(w)}(t,s) = \int_t^s h_{t0}^{(w)}(v)\exp\left(\mathbf{X}_p^{\intercal}\boldsymbol{\beta}_t^{(w)} + \mathbf{Z}_p^{\intercal}\boldsymbol{\phi}_t^{(w)} + \eta_t^{(w)}\xi_p(v) + \zeta^{(w)}\nu_p\right)\mathrm{S}_p(v)\mathrm{d}v \tag{3.7}$$

denote the subject-p-specific overall survival function across two competing-risk outcomes (i.e., the conditional probability that subject p has experienced neither of the competing terminal events till time t) and the cumulative incidence function for the wth competing-risk terminal event (i.e., the conditional probability that subject p experiences the wth competing-risk terminal event within the time interval $(t, s]$), respectively. Thus, the first component of the integrand in $\pi_p^{(w)}(t,s)$ can be addressed using (3.6), while the second component is estimated based on the posterior distribution of the model parameters from the training

dataset $\mathcal{D}_n$. A Monte Carlo estimate of $\pi_p^{(w)}(t, s)$ is obtained through a simulation scheme. The detailed algorithm for this process is provided in Appendix C.2, with a summarized procedure outlined in Table C.1 of Appendix C.

3.3.2 Predictive Accuracy Measures

To assess the predictive accuracy of the proposed BM-JM, we used two widely recognized metrics: the area under the receiver operating characteristic curve (AUC) and the expected Brier score (BS). AUC captures the model's overall ability to distinguish between different outcomes, while BS measures the average discrepancy between predicted risks and observed event probabilities. Given the dynamic nature of predictions in a competing risks framework, we utilized an adapted version of these measures to evaluate the BM-JM's performance across varying prediction time windows. To handle right-censored data, as seen in our CRIC study, we applied the Inverse Probability of Censoring Weighting (IPCW) method to estimate the dynamic AUC and BS, denoted as $\widehat{\text{AUC}}^{(w)}(t, s)$ and $\widehat{\text{BS}}^{(w)}(t, s)$ for the w-th competing risk. This approach accounts for censored observations, ensuring reliable performance evaluation. Detailed definitions, derivations, and implementation of these dynamic AUC and BS in a competing risks context, along with their computation under the BM-JM framework, are presented in Appendix C.3.

R codes and documentation for fitting BM-JM and predicting dynamically updated survival probabilities for future subjects, as well as predictive accuracy measures, are made publicly available at https://github.com/qqian-biostat/BM-JM. Due to the absence of closed-form solutions for the posterior distributions, the proposed model was fitted using the Bayesian software JAGS (version 4.3.0) via the `rjags` package (with more details deferred to Appendix C.1).

3.4 Application to the CRIC Study

3.4.1 Study Cohort

The CRIC study, initiated by NIDDK in 2001, is an ongoing multicenter prospective longitudinal cohort study that enrolls adults with mild to moderate CKD. After exclusion of patients with missing laboratory data (7.4%) and with CV hospital admission and death dates within one-day apart (0.2%), the final analysis cohort included 5,194 patients. Comprehensive baseline characteristics and pertinent risk factors of the analysis cohort are summarized in Table C.2. As outlined in Section 3.2.1, our approach segregates covariates considered into primary ($\mathbf{Z}_i$) and secondary ($\mathbf{X}_i$) risk factors. Primary risk factors include diabetes, hypertension, and history of CVD. Secondary risk factors include demographic variables (age, sex, race), behavioral factors (current smoker, body mass index (BMI)), cardiovascular parameters (systolic and diastolic blood pressure (SBP/DBP), high blood pressure (blood pressure (BP) $>$ 130/80 mmHg), angle-brachial index (ABI)), laboratory measures (hemoglobin A1c (HbA1c), calcium, creatinine (Roche adjusted), glucose), and use of angiotensin-converting enzyme inhibitor (ACEI) or angiotensin receptor blocker (ARB) therapy. Within the analysis cohort, age of patients ranged between 21 and 79 years (mean 59.6, standard deviation (SD) 10.7), where 43.4% of patients were females. Approximately 12.5% of patients were current smokers. In addition, 51.1%, 86.5%, and 27.7% of patients had a history of diabetes, hypertension, and CVD, respectively. The median follow-up duration extended to 11.8 years, with a maximum follow-up period of approximately 16.9 years. Three fourths of the analysis cohort had survived beyond five years.

Our analysis focused on the interdependent outcomes of longitudinal eGFR progression, recurrent hospitalizations due to CV events, and competing-risk terminal events defined as kidney failure and death in the CKD population. The mean baseline eGFR was at 48.4 ml/min/1.73m^2 (SD 15.6, median 48.2, first-third quartile (Q1-Q3) 37.1-58.8). The longitudinal eGFR trajectories of 200 randomly selected patients, given in Figure 3.1 (A), illustrate the heterogeneity in kidney function in the analysis cohort throughout follow-up. CV events

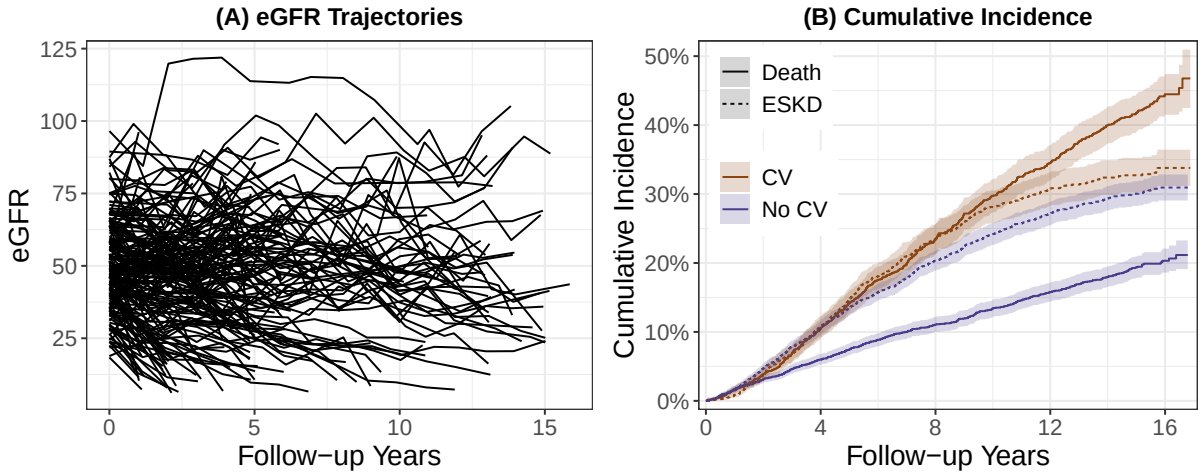

Figure 3.1: (A) Longitudinal estimated glomerular filtration rate (eGFR) trajectories for 200 randomly selected individuals, and (B) Cumulative incidence curves for ESKD and death from the analysis cohort patients with/out CV events.

considered included congestive heart failure, myocardial infarction, cerebrovascular accident, peripheral arterial disease and atrial fibrillation. In the recurrent event submodel, we focus on the first three hospitalizations due to CV events, typically of most clinical relevance. Within the analysis cohort, among the 1,582 subjects with one or more CV events, about 80% of patients experienced 1 (787), 2 (311), or 3 (161) CV events. The event rates for the two competing terminal events, ESKD and death, were comparable, at 25.53% and 20.91%, respectively. Figure 3.1 (B) shows the cumulative incidence of these events among CKD patients, stratified by the presence or absence of CV events. Notably, patients with CV events were more likely to progress to death, while those without CV events had a higher probability of progressing to ESKD. Furthermore, the progression to ESKD occurred at similar rates regardless of CV event status. However, the likelihood of death was significantly higher among CKD patients who experienced CV events, emphasizing the strong association between recurrent CV events and the risk of mortality.

3.4.2 Risk Factors and Interrelationships among Outcomes

The proposed BM-JM modeling framework was applied to the CRIC analysis cohort. To evaluate the predictive accuracy of the proposed dynamic prediction procedure, an inde-

pendent testing dataset of size $n_p = 50$ was randomly selected from the analysis cohort and excluded from the model fitting process. Results from the BM-JM are presented in Table 3.1 and Table C.3. Additional details on model fitting, including the specific choices of tuning and hyperparameters, the number of burn-in iterations, and the detailed prior distributions for model parameters, are provided in Appendix C.1.

Table 3.1: Results from the proposed BM-JM of (A) longitudinal estimated glomerular filtration rate (eGFR), (B) recurrent cardiovascular (CV) events, (C1 and C2) competing-risk terminal events (i.e., (C1) for ESKD and (C2) for death). Effect sizes (estimates and hazard ratios (HRs)) are given for one unit change in covariates.

	(A) Longitudinal eGFR	**(B) Recurrent CV Events**	**(C1) Terminal ESKD Event**	**(C2) Terminal Death Event**
Variable	Estimate (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Intercept	52.085 (51.323, 52.834)*	-	-	-
Time, γ	−1.505 (−1.592, −1.418)*	-	-	-
		[HTML]ECF0F1		
Age (years)	−0.304 (−0.327, −0.282)*	1.035 (1.028, 1.041)*	0.965 (0.957, 0.971)*	1.078 (1.060, 1.101)*
Non-Hispanic Black*	6.320 (5.836, 6.769)*	0.902 (0.807, 1.008)	1.313 (1.103, 1.564)*	0.930 (0.716, 1.197)
Other*	0.870 (0.233, 1.492)*	0.840 (0.715, 0.980)*	1.127 (0.923, 1.385)	1.037 (0.730, 1.484)
Female	−9.961 (−10.404, −9.497)*	0.890 (0.795, 0.998)*	0.454 (0.381, 0.544)*	0.478 (0.348, 0.633)*
		[HTML]ECF0F1		
Smoker	−0.423 (−1.093, 0.279)	1.391 (1.201, 1.611)*	1.188 (0.973, 1.441)	2.995 (2.100, 4.486)*
Body Mass Index	0.018 (−0.012, 0.048)	1.016 (1.009, 1.023)*	0.993 (0.983, 1.003)	1.016 (1.002, 1.031)*
		[HTML]ECF0F1		
ABI < 0.9	−0.135 (−0.729, 0.450)	1.371 (1.214, 1.550)*	1.089 (0.903, 1.309)	2.226 (1.644, 3.077)*
ACEI/ARB use	−1.179 (−1.686, −0.709)*	0.948 (0.845, 1.066)	1.047 (0.891, 1.229)	0.796 (0.608, 1.023)
Glucose (mg/dL)	0.007 (0.001, 0.012)*	1.000 (0.999, 1.002)	1.000 (0.998, 1.001)	0.999 (0.996, 1.002)
Hemoglobin A1c (%)	−0.237 (−0.439, −0.041)*	1.125 (1.078, 1.176)*	1.073 (1.013, 1.134)*	1.266 (1.137, 1.426)*
Hemoglobin (g/dL)	0.164 (0.029, 0.296)*	0.952 (0.920, 0.985)*	0.922 (0.880, 0.968)*	0.858 (0.791, 0.924)*
Calcium (mg/dL)	1.014 (0.577, 1.464)*	0.828 (0.746, 0.919)*	0.791 (0.685, 0.914)*	0.925 (0.736, 1.161)
Creatinine (mg/dL)	−25.497 (−25.941, −25.048)*	1.416 (1.274, 1.575)*	0.499 (0.419, 0.589)*	1.305 (0.954, 1.815)
		[HTML]ECF0F1		
High BP	−0.547 (−1.124, −0.028)*	0.997 (0.865, 1.156)	1.001 (0.810, 1.235)	0.857 (0.611, 1.198)
SBP (mmHg)	−0.032 (−0.049, −0.017)*	1.003 (1.000, 1.007)	1.005 (1.000, 1.010)	1.014 (1.005, 1.022)*
DPB (mmHg)	0.012 (−0.011, 0.032)	0.997 (0.992, 1.002)	0.999 (0.992, 1.006)	0.997 (0.985, 1.009)
		[HTML]ECF0F1		
Diabetes	−0.087 (−0.582, 0.408)	-	1.226 (1.014, 1.481)*	1.255 (0.939, 1.679)
Diabetes - ϕ_{11}	-	1.158 (1.003, 1.336)*	-	-
Diabetes - ϕ_{12}	-	0.961 (0.799, 1.162)	-	-
Diabetes - ϕ_{13}	-	0.787 (0.634, 1.018)	-	-
		[HTML]ECF0F1		
Hypertension	−1.588 (−2.262, −0.928)*	-	1.230 (0.945, 1.623)	1.380 (0.960, 1.998)
Hypertension - ϕ_{21}	-	1.618 (1.311, 2.042)*	-	-
Hypertension - ϕ_{22}	-	1.149 (0.838, 1.600)	-	-
Hypertension - ϕ_{23}	-	1.309 (0.874, 2.014)	-	-
		[HTML]ECF0F1		
CVD	−0.566 (−1.006, −0.126)*	[HTML]ECF0F1 -	1.290 (1.097, 1.510)*	3.065 (2.331, 4.233)*
CVD - ϕ_{31}	-	2.654 (2.349, 3.001)*	-	-
CVD - ϕ_{32}	-	1.718 (1.445, 2.028)*	-	-
CVD - ϕ_{33}	-	1.632 (1.315, 2.028)*	-	-
α_1	-	5.755 (3.885, 8.390)*	-	-
α_2	-	1.567 (0.926, 2.656)	-	-

1. $*$ 95% credible interval (CI) does not include estimate of 0 or hazard ratio (HR) of 1;
2. $\star$ Reference group: Non-Hispanic White;
3. Angiotensin-converting enzyme inhibitor(ACEI); angiotensin receptor blocker (ARB); blood pressure (BP); cardiovascular disease (CVD); hemoglobin A1c (HbA1c); systolic and diastolic BP (SBP, DBP)

As outlined in Table 3.1 (A), the longitudinal eGFR exhibited a yearly decline of -1.51 ml/min/1.73m^2 (95% credible interval (CI): -1.59 to -1.42) per year. Two primary risk factors were significantly associated with lower eGFR: hypertension (-1.59, CI: -2.26 to -0.93) and a history of CVD (-0.57, CI: -1.01 to -0.13). Additional factors associated with lower eGFR included being female (-9.96), ACEI/ARB therapy (-1.18), high BP (-0.55), older age (-0.30), and higher levels of creatinine (-25.50), HbA1c (-0.24) and SBP (-0.03). Conversely, factors associated with higher eGFR included being black (6.32), belonging to a race other than white or black (0.87), and higher levels of calcium (1.01), hemoglobin (0.16), and glucose (0.01). Notably, diabetes, the third primary risk factor, was not found to be significantly associated with eGFR after adjusting for relevant laboratory markers (e.g., HbA1c). Insights from Table C.3 (B) further reveal that subject-specific variation accounted for a substantial proportion of total variation in eGFR. Additionally, a positive correlation ($\rho_b = 0.31$) was observed between the random effects for the intercept and slope.

Next, we explored the risk factors associated with recurrent CV events, as outlined in Table 3.1 (B). Among the secondary risk factors, being a smoker (hazard ratio (HR) 1.39), having an ABI < 0.9 (HR 1.37), older age (HR 1.04), and higher levels of BMI (HR 1.02), HbA1c (HR 1.13), and creatinine (HR 1.42) were significantly associated with an increased hazard of recurrent CV events. Conversely, belonging to a race other than white or black (HR 0.84), being female (HR 0.89), and having higher levels of hemoglobin (HR 0.95) and calcium (HR 0.83) were associated with a decreased hazard of CV events. The event-varying effects of primary risk factors on the 1st, 2nd and 3rd CV events, as well as the event effects (i.e., the effect of prior CV events on future CV events), were also studied. As shown in Table 3.1 (B), patients with a history of CVD had the highest HR for the first CV event (HR 2.65, CI: 2.35 to 3.00) compared to those without CVD. Although the hazard remained significantly elevated for the second (HR 1.72, CI: 1.45 to 2.03) and the third (HR 1.63, CI: 1.32 to 2.03) CV events, it was reduced by more than half compared to the risk for the first CV event. Patients with hypertension exhibited the second-highest HR for the first CV event (HR 1.62, CI: 1.31 to 2.04), while patients with diabetes showed only a 16% increased hazard

for the first CV event (HR 1.16, CI: 1.00 to 1.34), potentially due to an indirect effect of diabetes mediated through HbA1c control. Notably, neither hypertension nor diabetes were significantly associated with subsequent CV events. Regarding the CV event effects ($\widehat{\alpha}$'s), the hazard of experiencing the next CV event for a CKD patient with one prior CV event was substantial compared to a patient with no prior events (HR 5.76, CI: 3.89 to 8.39). However, for a patient with two past CV events, the hazard of experiencing the next (third) CV event was drastically reduced and not statistically significant (HR 1.57, CI: 0.93 to 2.66).

For the risk factors influencing competing-risk terminal events (ESKD and death) (Tables 3.1 (C1) and (C2)), a history of CVD was associated with higher hazards for both events, with a particularly elevated hazard for death (HR 3.07, CI: 2.33 to 4.23). Diabetes was associated with an increased hazard of ESKD only (HR 1.23, CI: 1.01 to 1.48), while hypertension was not significantly associated with either event. Elevated levels of HbA1c were linked to an increased hazard for both competing risks, whereas being female and having higher hemoglobin levels were associated with reduced hazard for both ESKD and death. Interestingly, age had opposing effects on the two competing risks, i.e., higher age was associated with a lower hazard for ESKD but a higher hazard for death. Additionally, some predictors had impacts only on one competing risk. For instance, being black (HR 1.31) was associated with a higher hazard of ESKD, while higher levels of hemoglobin (HR 0.92), calcium (HR 0.79) and creatinine (HR 0.50) were associated with reduced hazards of ESKD. In contrast, smoking (HR 3.00) and ABI < 0.9 (HR 2.23) were significantly associated with an increased hazard of death.

Regarding the interrelationships among multivariate outcomes (shown in Appendix C Table C.3 (A)), individual variability explained by the random slope in longitudinal eGFR trajectories was significantly associated with recurrent CV events ($\eta_{r1} = -0.16$, CI: -0.19 to -0.13). This strong association aligns with expectations, as declining kidney function is well-documented to correlate with CV events in CKD patients (Go et al., 2004). In contrast, variability in the random intercept was not significantly associated with CV events ($\eta_{r0} = 0.004$, CI: -0.006 to 0.014). As highlighted in Section 3.1 and 3.2.1, a key innovation of the proposed BM-JM is its capacity to directly quantify associations between the longitudinal

biomarker and competing-risk terminal events via $\eta_t^{(w)}$. Lower eGFR levels were strongly associated with ESKD ($\eta_t^{(1)} = -0.19$, CI: -0.20 to -0.18), while the direct effect of lower eGFR levels on death was much smaller ($\eta_t^{(2)} = -0.02$, CI: -0.03 to -0.01). This pattern is expected, as declining kidney function directly contributes to kidney failure, whereas the causes of death in CKD patients are more diverse and complex. Finally, recurrent CV events demonstrated a much stronger positive association with death ($\zeta^{(2)} = 2.82$, CI: 2.06 to 3.69) compared to their association with ESKD ($\zeta^{(1)} = 0.29$, CI: 0.07 to 0.51), which aligns with the higher probability of progression to death observed among CKD patients with CV events observed in Figure 3.1 (B).

3.4.3 Comparison with Simpler Models

We compared the proposed BM-JM with other well-established models that overlook the joint outcomes structure. Specifically, our proposed BM-JM consists of three components: the longitudinal, recurrent and competing-risk terminal components. We decompose these components into three separate marginal models, each addressing a single outcome: the longitudinal model (L-M), the recurrent model (R-M) and the competing-risk terminal model (T-M). Additionally, we combine two of these components to form simplified joint models. Since one of our primary objectives is to make dynamic predictions for competing risks, we focus on two simplified joint models: (1) the joint model of longitudinal measurements and competing-risk terminal events (referred to as LT-JM), which excludes the recurrent component, and (2) the joint model of recurrent and competing-risk terminal events (referred to as RT-JM), which excludes the longitudinal component. All these models were applied to the CRIC study analysis cohort. Further details about these comparative models are provided in Appendix C.4, with results from three marginal models summarized in Appendix C Table C.4 (A-B-C) for the longitudinal eGFR (L-M), recurrent CV events (R-M), and competing-risk terminal events (T-M), respectively. Additionally, the results for the LT-JM and RT-JM fits are presented in Tables C.5 and C.6, respectively.

Comparison with marginal models: For the longitudinal eGFR, the results from the marginal model (L-M) (see Table C.4 (A)) were largely consistent with those of the pro-

posed BM-JM. The effects of all covariates on eGFR trajectories, including their inference, were comparable between the two models. Notably, the primary covariates (diabetes, hypertension, and history of CVD) exhibited similar effects on eGFR across both models. However, the decline in kidney function over time was slightly attenuated in the marginal model (L-M -1.33 vs. BM-JM -1.51). For the recurrent CV events, while the overall conclusions for the primary covariates were similar between the marginal R-M (see Table C.4 (B)) and the BM-JM analyses, in terms of effect sizes and inference, some notable differences regarding the secondary covariates emerged. For instance, the effect of creatinine on recurrent CV events was smaller in the marginal R-M analysis (R-M HR 1.12 vs. BM-JM HR 1.42).

For the competing risks of ESKD and death, the marginal T-M analysis (see Table C.4 (C1 and C2)) showed similar inference based on CIs for the primary covariates compared to the BM-JM. However, there were notable differences in effect estimates. For instance, the marginal T-M analysis yielded higher effect estimates for diabetes and history of CVD on ESKD but lower effect estimates for death compared to the BM-JM. Additionally, the T-M analysis estimated a stronger effect of smoking on ESKD (T-M HR 1.64 vs. BM-JM HR 1.19) but a weaker effect on death (T-M HR 1.80 vs. BM-JM HR 3.00). Notably, the marginal T-M analysis produced conflicting results for creatinine and female gender on ESKD compared to the BM-JM, i.e., the T-M estimated an HR of 11.82 for creatinine and 1.31 for female gender on ESKD, while the BM-JM estimated HRs of 0.50 and 0.45, respectively. Furthermore, while the effect of creatinine was statistically significant for death in the T-M analysis, it was not significant in the BM-JM.

Comparison with simpler joint models: For the LT-JM model fits (see Table C.5), the results for the longitudinal eGFR and ESKD terminal submodels were generally consistent with those obtained from the BM-JM. However, notable differences were observed in the death terminal submodel. Specifically, the effect estimates for being a smoker, having an ABI < 0.9, and a history of CVD on mortality were significantly attenuated in the LT-JM. This suggests that excluding the recurrent component impacts the death submodel's estimates, as expected, due to the stronger association of recurrent CV events with mortality.

In comparison, the differences between the RT-JM (see Table C.6) and the BM-JM anal-

yses were more pronounced. While the overall effects of covariates on recurrent CV events were similar across the two models, the estimates and inferences for the competing-risk components showed significant disparities. Except for the effect of a history of CVD on mortality, the effects of primary covariates on ESKD and death were markedly amplified in the RT-JM compared to the BM-JM. Substantial differences were particularly evident in the ESKD terminal submodel, which is more closely associated with longitudinal eGFR measurements than mortality. For instance, the RT-JM estimated an HR of 9.07 for creatinine and 1.21 for female gender, compared to HRs of 0.50 and 0.45, respectively, in the BM-JM. Additionally, the effect of creatinine on mortality was statistically significant in the RT-JM but not in the BM-JM. Regarding the correlation and covariance estimates, the variance for the frailty term ν_i and the correlation parameters $\zeta^{(w)}$ were significantly larger in the RT-JM. This is likely attributable to the omission of variability from the longitudinal component in each submodel of the RT-JM. As discussed in the next section, the changes in model fits using the LT-JM and RT-JM led to poorer dynamic prediction performance for both ESKD and death compared to the BM-JM. These performance gaps were particularly pronounced in the AUC and BS assessments for mortality predictions, where both simplified models demonstrated significantly lower accuracy.

3.4.4 Dynamic Prediction Results

Next, posterior samples from the BM-JM fit were used to estimate the cumulative incidence functions for ESKD and death, given a new patient's baseline characteristics, available eGFR measurements, and recurrent CV event profiles, as outlined in Section 3.3 and Table C.1 (of Appendix C). Appendix C Figures C.9 and C.10 (C-J) present the median of 200 Monte Carlo realizations of $\pi_p^{(w)}(s,t)$, along with 95% pointwise CIs based on the percentiles of $\pi_p^{(w)}(s,t)$ over the Monte Carlo samples. These predictions are displayed for varying prediction times $t = 0, 2, 4, 6$ up to 14 years of follow-up for two patient pairs.

The first pair, Patients 23 and 24, shared nearly identical baseline characteristics: both were 58-year-old male CKD patients of other race with diabetes, hypertension, and a history

of CVD at enrollment. Additionally, both experienced three recurrent CV events during follow-up (Figure C.9 (B)). However, their eGFR trajectories diverged significantly: Patient 23 maintained a relatively stable and higher eGFR profile throughout follow-up, whereas Patient 24 exhibited a steady decline in eGFR starting two years into follow-up (Figure C.9 (A)). These differences in eGFR trajectories had a profound impact on their predicted risks (as shown in Figure C.9 (C-J)). Patient 24 had a probability exceeding 0.6 of developing ESKD by year 6, which rose to over 0.8 by year 14, consistent with the observed outcome of being diagnosed with ESKD at 8.61 years. In contrast, Patient 23, despite having a slightly higher predicted mortality risk, displayed a much lower probability of developing ESKD, aligning with the observed outcome of death at 12.69 years. These results highlight the critical role of eGFR trajectories in predicting ESKD risk and demonstrate how incorporating longitudinal profiles can refine individualized risk predictions.

The second pair, Patients 12 and 10, also shared similar baseline characteristics: both were approximately 70-year-old male CKD patients of other race with diabetes, hypertension, and a history of CVD at baseline. Their eGFR trajectories during follow-up were roughly similar (see Figure C.10 (A)), with Patient 12 exhibiting a unique dip in eGFR values towards the end of follow-up, around 10 years. However, their recurrent CV event profiles differed significantly: Patient 12 experienced no recurrent CV events during follow-up, whereas Patient 10 underwent three successive CV events (Figure C.10 (B)). As shown in Figure C.10 (C-J), the ESKD risk predictions for these two patients were relatively similar, especially at $t = 6$ years. However, at earlier time points, when only baseline characteristics ($t = 0$) and eGFR trajectories up to two years ($t = 2$) were considered, Patient 12 exhibited a higher predicted risk of ESKD. In contrast, their mortality risk predictions diverged substantially: Patient 10's predicted mortality risk escalated to over 0.6 by year 14, whereas Patient 12's predicted mortality risk remained around 0.1. These predictions are consistent with the observed outcomes: Patient 12 progressed to ESKD towards the end of follow-up at year 10.77, while Patient 10 died at year 7.78, shortly after his last CV event. The inclusion of recurrent CV event data heavily influenced the differences in survival risk predictions for this pair. Additionally, the increased variability in predictions for Patient 10

at later years could be attributed to the reduced number of patients remaining in the study, leading to greater uncertainty in estimated hazards. This example, alongside the analysis of Patients 23 and 24, highlights the importance of integrating both longitudinal measurements and recurrent event data in prognostic modeling for CKD patients. Notably, these findings align with conclusions drawn from $\eta_t^{(w)}$ and $\zeta^{(w)}$, as well as observations in Figure 3.1 (B), underscoring the strong correlation between longitudinal eGFR trajectories and ESKD, and between recurrent CV events and mortality.

To further evaluate the predictive accuracy of the dynamic prediction results obtained from the BM-JM and compare them with those from the comparative joint models, LT-JM and RT-JM, dynamic AUC and BS values were calculated using the testing dataset ($n_p = 50$) which was set aside from model fitting. The results (summarized in Figure C.11) indicate that the overall prediction performance for ESKD surpasses that for mortality, as evidenced by higher AUC and lower BS values across all models in most time points. This finding aligns with expectations, as longitudinal eGFR trajectories are strong markers for kidney decline and progression to ESKD, whereas mortality in CKD patients is inherently more challenging to predict. For a fixed prediction time t, dynamic AUC values tend to decrease, and dynamic BS values tend to increase with a longer prediction time window Δt. This trend arises because even small differences between the true and estimated random effects are expected to translate to greater differences between the corresponding cumulative incidence functions as time progresses (i.e., for greater Δt). Notably, prediction accuracy does not necessarily improve as the prediction time t increases, despite the availability of additional longitudinal and recurrent data. This phenomenon reflects the interplay of two opposing factors: the accumulation of additional longitudinal and recurrent event information over time reduces the degree of shrinkage in the estimation of random effects. However, as the prediction time progresses, fewer patients remain in the study, leading to increased variability in the estimated hazard functions. These opposing forces may result in prediction accuracy remaining steady as t increases.

Importantly, the proposed BM-JM consistently outperforms both LT-JM and RT-JM, as evidenced by higher dynamic AUC and lower dynamic BS values for both ESKD and mor-

tality risk predictions. This performance gap widens further as prediction time t increases, due to the increasing availability of recurrent/longitudinal data. While RT-JM, which incorporates recurrent CV events into the joint modeling, outperforms LT-JM, especially for mortality predictions at earlier prediction times (i.e., $t = 0$), LT-JM begins to surpass RT-JM as prediction time progresses (starting from $t = 2$). This shift is due to the growing availability of longitudinal measurements data but decreasing availability of recurrent CV events information (as fewer patients remain in the study over time). These findings highlight the complementary roles of recurrent and longitudinal data in joint modeling frameworks, particularly in improving the accuracy of competing-risk terminal outcome predictions.

3.5 Simulation Studies

Simulation studies were conducted to evaluate the performance of the proposed estimation and dynamic prediction procedures and to compare the proposed BM-JM with three marginal models (L-M, R-M, and T-M) and two simplified joint models (LT-JM and RT-JM). The performance of the proposed estimators was assessed using bias, mean-squared error (MSE), and coverage probabilities (CP) of 95% credible intervals (CIs). Simulation results, based on 200 Monte Carlo runs, are summarized in Table C.7 and Table C.8 of Appendix C, with detailed simulation setups provided in Appendix C.5. The results in Table C.7 demonstrate that the BM-JM achieves low bias and MSE values across all parameters in the longitudinal, recurrent, and competing-risk terminal submodels, with CPs consistently close to the expected 95%, underscoring the robustness of the proposed estimation and inference procedure. In contrast, the three marginal models (L-M, R-M, and T-M with results summarized in Table C.8) and the two simplified joint models (LT-JM and RT-JM with results summarized in Table C.7) exhibited larger biases, particularly in the competing-risk terminal submodels. Additionally, the efficacy of inference using CIs for the competing-risk terminal submodels was notably compromised in all comparative models. These findings emphasize the necessity of caution when interpreting results derived from simpler models.

The proposed dynamic prediction algorithm was also evaluated in simulation studies,

with results based on 200 testing datasets reported in Appendix C Table C.9. The Dynamic AUC and BS reported for BM-JM demonstrate commendable predictive accuracy across different settings, with all AUC values exceeding 0.74 and BS values falling below 0.20. Additionally, with a fixed follow-up time t and an increasing prediction time window Δt, dynamic AUC values tend to decrease, while dynamic BS values tend to increase, similar to findings reported in predictions of the CRIC data. Furthermore, prediction accuracy does not necessarily improve with increasing time t, despite the availability of more longitudinal measurements and recurrent CV events information. As explained in the CRIC data analysis results, this phenomenon can be attributed to the fact that while there is a reduction in the degree of shrinkage in the estimation of the random effects over time, there is also greater variability in the estimated hazard functions as the prediction time progresses, due to fewer subjects remaining in the study. Finally, similar to the CRIC study findings, the LT-JM and RT-JM models consistently exhibit lower dynamic AUC values and higher dynamic BS values compared to the proposed BM-JM across all settings. The performance gains of the BM-JM are particularly pronounced in later periods, where additional information from recurrent events and longitudinal measurements becomes available. These results underscore the importance of incorporating both longitudinal biomarker and recurrent event data into joint modeling frameworks to enhance predictive accuracy.

3.6 Discussion

Contributions of this work are multifold. First, the proposed BM-JM addresses the methodological gap in multivariate joint modeling of longitudinal trajectories, recurrent events and competing-risk terminal events, while also providing an efficient dynamic prediction framework for competing-risk terminal outcomes utilizing the combined history of longitudinal measurements and recurrent events. A unique feature of the proposed BM-JM is the explicit modeling of the effects of longitudinal biomarker in the competing-risk survival submodels, allowing for direct quantification and interpretation of the predictive value of the longitudinal measurements on competing-risk terminal outcomes. In addition, the proposed framework

allows for prediction of each competing risk individually, leveraging complete longitudinal profile and recurrent events information over time, enhancing predictive accuracy of BM-JM significantly. Dynamic prediction framework provides clinicians with evidence-based insights into the prognostic implications associated with changes in a patient's disease condition.

Additionally, our work, directly motivated by the primary goals of the CRIC study, offers unique findings on CKD progression. The analysis includes the study of the complex network of risk factors influencing key interrelated multivariate outcomes of kidney disease progression, recurrent hospitalizations due to CV events and the eventual competing-risk terminal events of ESKD and death. Dynamic predictions of ESKD and death can serve as an early warning system, offering clinicians valuable lead time to plan and implement interventions effectively. Findings highlight the strong associations between longitudinal eGFR trajectories and ESKD, and recurrent CV events and mortality and hence the importance of considering both longitudinal measurements and recurrent events data in prognostic modeling for CKD patients, as they provide complementary insights into disease progression and outcomes. Finally, the flexibility of our proposed estimation and prediction procedure, developed within a competing risk framework in a Bayesian paradigm, allows for its adaptation to complex scenarios. While our illustration concentrates on one longitudinal, one recurrent and two competing-risk terminal processes, the proposed framework can easily be extended to more intricate situations, including multiple longitudinal biomarkers and more than two competing-risk terminal processes.

While the current joint model employs a linear structure for the longitudinal biomarker, it can be generalized to account for non-linear effects of time if the data suggest such patterns, including higher-order polynomials (e.g., t^2), splines, or other non-linear functions of time. Additionally, exploring alternative noise distributions, such as the extreme value distribution, could provide further insights into model behavior. Although these extensions are beyond the primary scope of this study, they offer promising directions for future research. Moreover, variational Bayes inference (VBI) has gained significant attention in the context of complex machine learning and statistical models for its ability to efficiently approximate posterior distributions. While the complexity of our current model can be adequately handled using

the standard MCMC approach, we acknowledge the potential of VBI as a valuable future direction, especially in scenarios with increased computational demands or further model extensions.

Appendix A

Supporting Information for Chapter 1

A.1 Derivation of the Form of the Multivariate Eigenfunctions

Let $\boldsymbol{\psi}^{(j)}(t) = \{\psi_1^{(j)}(t), \cdots, \psi_{M^+}^{(j)}(t)\}$. By the orthonormality of the multivariate eigenfunctions and the CAR model assumed on the multivariate PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\top \sim N(\mathbf{0}, \alpha_\ell \Sigma_w)$, where $\Sigma_w = (D - \nu W)^{-1}$ and $\circ$ denotes elementwise multiplication,

$$
\begin{aligned}
&\sum_{j'=1}^{J} \int_{\mathcal{T}} \mathrm{Cov}(X_i^{(j)}(t), X_{i'}^{(j')}(s)) \boldsymbol{\psi}^{(j')}(s) \mathrm{d}s \\
=&\sum_{j'=1}^{J} \sum_{\ell=1}^{M^+} \int_{\mathcal{T}} \mathrm{Cov}(\rho_{i\ell}, \rho_{i'\ell}) \psi_\ell^{(j)}(t) \psi_\ell^{(j')}(s) \boldsymbol{\psi}^{(j')}(s) \mathrm{d}s \\
=&\left\{ \mathrm{Cov}(\rho_{i1}, \rho_{i'1}) \psi_1^{(j)}(t), \cdots, \mathrm{Cov}(\rho_{iM^+}, \rho_{i'M^+}) \psi_{M^+}^{(j)}(t) \right\} \\
=&(\Sigma_w)_{ii'} \left\{ \alpha_1 \psi_1^{(j)}(t), \cdots, \alpha_{M^+} \psi_{M^+}^{(j)}(t) \right\} \\
=&(\Sigma_w)_{ii'} \overbrace{(\alpha_1 \cdots, \alpha_{M^+})}^{\equiv \boldsymbol{\alpha}} \circ \overbrace{\left(\psi_1^{(j)}(t), \cdots, \psi_{M^+}^{(j)}(t) \right)}^{= \boldsymbol{\psi}^{(j)}(t)} \\
=&(\Sigma_w)_{ii'} \boldsymbol{\alpha} \circ \boldsymbol{\psi}^{(j)}(t).
\end{aligned} \tag{A.1}
$$

Assuming the outcome in each dimension have a finite K-L representation, i.e., $X_i^{(j)}(t) = \sum_{m=1}^{M_j} \xi_{im}^{(j)} \phi_m^{(j)}(t)$, $j = 1, \ldots, J$, and letting $\Sigma_{b_{m'm''}}^{(jj')}$ denote the (m', m'')th element of the (j, j')

block of Σ_b, (A.1) also equals

$$\begin{aligned}
&\sum_{j'=1}^{J}\int_{\mathcal{T}}\mathrm{Cov}\left(X_i^{(j)}(t),X_{i'}^{(j')}(s)\right)\boldsymbol{\psi}^{(j')}(s)\mathrm{d}s\\
=\;&\sum_{j'=1}^{J}\int_{\mathcal{T}}\mathrm{Cov}\left(\sum_{m'=1}^{M_j}\xi_{im'}^{(j)}\phi_{m'}^{(j)}(t),\sum_{m''=1}^{M_{j'}}\xi_{i'm''}^{(j')}\phi_{m''}^{(j')}(s)\right)\boldsymbol{\psi}^{(j')}(s)\mathrm{d}s\\
=\;&\sum_{j'=1}^{J}\sum_{m'=1}^{M_j}\sum_{m''=1}^{M_{j'}}\overbrace{\mathrm{Cov}\left(\xi_{im'}^{(j)},\xi_{i'm''}^{(j')}\right)}^{=\Sigma_{b_{m'm''}}^{(jj')}\times(\Sigma_w)_{ii'}}\phi_{m'}^{(j)}(t)\overbrace{\int_{\mathcal{T}}\phi_{m''}^{(j')}(s)\boldsymbol{\psi}^{(j')}(s)\mathrm{d}s}^{\equiv\boldsymbol{c}_{m''}^{(j')}}\\
=\;&(\Sigma_w)_{ii'}\sum_{j'=1}^{J}\sum_{m'=1}^{M_j}\sum_{m''=1}^{M_{j'}}\Sigma_{b_{m'm''}}^{(jj')}\phi_{m'}^{(j)}(t)\boldsymbol{c}_{m''}^{(j')}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&(\Sigma_w)_{ii'}\sum_{j'=1}^{J}\sum_{m'=1}^{M_j}\sum_{m''=1}^{M_{j'}}\Sigma_{b_{m'm''}}^{(jj')}\phi_{m'}^{(j)}(t)\boldsymbol{c}_{m''}^{(j')}=(\Sigma_w)_{ii'}\boldsymbol{\alpha}\circ\boldsymbol{\psi}^{(j)}(t)\\
\Rightarrow\;&\sum_{j'=1}^{J}\sum_{m'=1}^{M_j}\sum_{m''=1}^{M_{j'}}\Sigma_{b_{m'm''}}^{(jj')}\phi_{m'}^{(j)}(t)\boldsymbol{c}_{m''}^{(j')}=\boldsymbol{\alpha}\circ\boldsymbol{\psi}^{(j)}(t) \qquad \text{(A.2)}
\end{aligned}$$

for $i,i'=1,\ldots,n$. Multiplying both sides of (A.2) by $\phi_m^{(j)}(t)$ and integrating over $\mathcal{T}$,

$$\begin{aligned}
&\int_{\mathcal{T}}\phi_m^{(j)}(t)\sum_{j'=1}^{J}\sum_{m'=1}^{M_j}\sum_{m''=1}^{M_{j'}}\Sigma_{b_{m'm''}}^{(jj')}\phi_{m'}^{(j)}(t)\boldsymbol{c}_{m''}^{(j')}\mathrm{d}t=\int_{\mathcal{T}}\phi_m^{(j)}(t)\boldsymbol{\alpha}\circ\boldsymbol{\psi}^{(j)}(t)\mathrm{d}t\\
\Rightarrow\;&\sum_{j'=1}^{J}\sum_{m''=1}^{M_{j'}}\overbrace{\int_{\mathcal{T}}\sum_{m'=1}^{M_j}\Sigma_{b_{m'm''}}^{(jj')}\phi_m^{(j)}(t)\phi_{m'}^{(j)}(t)\mathrm{d}t}^{\Sigma_{b_{mm''}}^{(jj')}}\boldsymbol{c}_{m''}^{(j')}=\boldsymbol{\alpha}\circ\overbrace{\int_{\mathcal{T}}\phi_m^{(j)}(t)\boldsymbol{\psi}^{(j)}(t)\mathrm{d}t}^{\boldsymbol{c}_m^{(j)}}\\
\Rightarrow\;&\sum_{j'=1}^{J}\sum_{m''=1}^{M_{j'}}\Sigma_{b_{mm''}}^{(jj')}\boldsymbol{c}_{m''}^{(j')}=\boldsymbol{\alpha}\circ\boldsymbol{c}_m^{(j)} \qquad \text{(A.3)}
\end{aligned}$$

for $m=1,\cdots,M_j$ due to orthonormality of $\phi_m^{(j)}(t)$. Since m and j are arbitrarily chosen,

this is equivalent to

$$\begin{pmatrix} \Sigma_b^{(11)} & \cdots & \Sigma_b^{(1J)} \\ \vdots & \ddots & \vdots \\ \Sigma_b^{(J1)} & \cdots & \Sigma_b^{(JJ)} \end{pmatrix} \begin{pmatrix} \mathbf{c}^{(1)} \\ \vdots \\ \mathbf{c}^{(J)} \end{pmatrix} = \boldsymbol{\alpha} \circ \begin{pmatrix} \mathbf{c}^{(1)} \\ \vdots \\ \mathbf{c}^{(J)} \end{pmatrix}. \tag{A.4}$$

with matrices $\Sigma_b^{(jj')} \in \mathbb{R}^{M_j \times M_{j'}}$ and $\mathbf{c}^{(j)} \in \mathbb{R}^{M_j}$. Since α_ℓ's are spatial variance parameters with $\alpha_\ell \geq 0$, (A.4) is an eigenequation for the symmetric positive semidefinite block matrix $\Sigma_b \in \mathbb{R}^{M^+ \times M^+}$. Let $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_L > 0$, $L \leq M^+$ be the non-zero eigenvalues of Σ_b and the elements $\psi_\ell^{(j)}(t)$ of the associated eigenfunctions $\psi_\ell(t)$ are parametrized by the (orthonormal) eigenvectors $\mathbf{c}_1, \ldots, \mathbf{c}_L$ associated with $\alpha_1, \ldots, \alpha_L$:

$$\psi_\ell^{(j)}(t) \overset{(A.2)}{=} \frac{1}{\alpha_\ell} \sum_{j'=1}^{J} \sum_{m'=1}^{M_j} \sum_{m''=1}^{M_{j'}} \Sigma_{b_{mm''}}^{(jj')} [\mathbf{c}_\ell]_{m''}^{(j')} \phi_{m'}^{(j)}(t) \overset{(A.3)}{=} \sum_{m'=1}^{M_j} [\mathbf{c}_\ell]_{m'}^{(j)} \phi_{m'}^{(j)}(t),$$

for $\ell = 1, \ldots, L$ and $j = 1, \ldots, J$.

A.2 Description of the USRDS Study Cohort

Our study cohort includes patients of age 18 years or older who transitioned to dialysis between January 1, 2005, and September 30, 2013. The observation period starts from day 91 of dialysis and patients are followed up for two years where the last date of follow-up is December 31, 2015. The main eligibility criteria are that (a) patients survive the first 90 days of dialysis and do not recover renal function or receive kidney transplantation and (b) that they have Medicare as their primary payer on day 91 of dialysis. In obtaining the multivariate outcome, we exclude facilities with less than 10 patients in any one month follow-up interval (17.4%) to obtain stable facility-specific hospitalization and mortality rates over time. In addition, HSAs are merged first to guarantee that they contain at least 5 facilities and for a second time to guarantee that the total number of months with zero mortality rate does not exceed one-third of the two year follow-up. Both merging algorithms have four steps: Step 1) start with HSAs as regions; Step 2) among regions with the lowest number of facilities

(or with greater than eight months of zero mortality), randomly select one region (call it R_1); Step 3) among neighbors of the selected region R_1, randomly select a neighbor, among those with the lowest number of facilities (or among those with the largest number of months with zero mortality rate), to merge with region R_1; Step 4) repeat Steps 2 and 3 until the lowest number of facilities within a region is 5 (or the largest number of months with zero mortality rate during the two-year follow up within a region is eight). The 719 total HSAs are reduced to 402 after the first merging algorithm and to 367 after the second, containing 5497 facilities. The medium number of facilities in each HSA is 9 (first-third quartile: 7-15), while the medium number of patients in each facility per month is 57 (first-third quartile: 36-87). Note that the majority (61.6%) of the resulting 367 regions (merged HSAs) are unaffected, while 15.0%, 8.4%, and 15.0% of the merged regions consist of merging 2, 3, and > 3 original HSAs, respectively.

A.3 Description of Three Comparison Models

We compare the MST-FPCA fits to fits from three alternate methods of modeling multivariate spatiotemporal data. The first comparative approach (referred to as multivariate FPCA (M-FPCA)) obtains the multivariate eigenfunctions and eigenvalues using the multivariate FPCA algorithm proposed by Happ and Greven (2018). Assuming that multivariate principal component (PC) scores are i.i.d. across regions, inference on the multivariate PC scores are also targeted within a Bayesian hierarchical modeling framework via MCMC. With i.i.d. multivariate PC scores, M-FPCA ignores the spatial correlations across regions in the data. The second comparative approach considered is called the Multivariate 'Besag-York-Mollié' model (MBYM) proposed by Boulieri et al. (2017). MBYM includes multivariate spatially structured and unstructured effects, as well as a temporal component to capture the dependencies among multivariate outcomes, across regions and time points. While a random walk of order one is utilized to induce temporal correlations, a multivariate intrinsic conditional autoregressive (ICAR) model is utilized to model spatial correlations, which leads to a discrete-time, discrete-space modeling for multivariate spatiotemporal data. As it utilizes

a hierarchical model with separate additive random effects for space and time, it induce a separable space-time covariance structure. MBYM is easy to fit via the built-in `mv.car` package in WinBUGS, hence is utilized in comparisons in the analysis of the USRDS data and in simulations. The last alternate method considered is to fit univariate spatiotemporal FPCA separately in each dimension, which we refer to as U-FPCA. Similar to Step 1 of the MST-FPCA estimation algorithm, the U-FPCA approach targets the univariate eigenfunctions via FPCA as the first step. Then in each dimension, a CAR structure is induced among PC scores to capture dependencies among regions. Conditional on the estimated univariate eigenfunctions, the U-FPCA approach models the univariate functional outcome also within a Bayesian hierarchical modeling framework via MCMC.

A.4 Simulation Set-up for Multivariate Response in Two Dimensions

In the first independent multivariate response set-up, the bivariate response is generated as $X_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}^{(j)} \psi_\ell^{(j)}(t) + \epsilon_i^{(j)}(t)$ for $i = 1, \ldots, n$, $j = 1, 2$ where the time grid $\{t_1, \ldots, t_k : k = 1, \ldots, T\}$ is taken to be $T = 24$ or $T = 50$ equidistant time points between 0 and 1, with $T = 24$ mimicking the 24 month follow-up in our data application and $T = 50$ representing medium to dense time grids. To generate the univariate eigenfunctions, we consider 11 Fourier basis functions on the interval $\mathcal{T} = [0, 2]$. While $\psi_\ell^{(1)}(t)$ is taken to be the 4th, 5th and 8th Fourier basis functions evaluated on $[0, 1]$ and $\psi_\ell^{(2)}(t)$ is set to equal the first, 5th and 8th basis functions evaluated on $[1, 2]$, shifted to the left by 1 and multiplied by a random sign. Utilizing the map of $n = 49$ states in the contiguous U.S. (including the District of Columbia), and the map of the $n = 367$ HSAs in our data application, the region-specific PC scores, $\boldsymbol{\rho}_\ell^{(j)} = (\rho_{1\ell}^{(j)}, \rho_{2\ell}^{(j)}, \ldots, \rho_{n\ell}^{(j)})^\top$, along each dimension, are generated independently from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell^{(j)}(D - \nu W)^{-1}$, where W is the adjacency matrix and D is the diagonal matrix with number of regions, as described in Section 1.2.1 of main paper. The variance parameters $\alpha_\ell^{(j)}$'s are set to $\alpha_1^{(1)} = \alpha_1^{(2)} = 2.0$, $\alpha_2^{(1)} = \alpha_2^{(2)} = 1.2$, $\alpha_3^{(1)} = \alpha_3^{(2)} = 0.4$ and the smoothing parameter

$\nu = 0.9$. The measurement errors, $\epsilon_i^{(j)}(t)$, are generated from $N(0, \sigma^2)$, where σ^2 is set to 0.02 and 0.5 corresponding low and high noise, respectively.

In the second, more general, correlated response set-up the bivariate functional data is generated from the MST-FPCA model $X_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}\psi_\ell^{(j)}(t) + \epsilon_i^{(j)}(t)$ for $i = 1, \ldots, n$ and $j = 1, 2$. Distinct from the first set-up, $\psi_\ell^{(1)}(t)$ is taken to be the 4th, 5th and 8th Fourier basis functions evaluated on $[0, 1]$ and $\psi_\ell^{(2)}(t)$ is set to equal 4th, 9th and 8th basis functions evaluated on $[1, 2]$, shifted to the left by 1 and multiplied by a random sign. The sign of $\psi_\ell^{(2)}(t)$ is selected to guarantee that the multivariate eigenfunctions are orthonormal on $[0, 1]$. In addition, the dependency of the multivariate response is induced by the shared PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\top$, generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$. The spatial variance parameters α_ℓ's are taken to be $\alpha_1 = 2.0$, $\alpha_2 = 1.2$ and $\alpha_3 = 0.4$, ν is set to 0.9, the measurement errors are generated similar to the first set-up. Note that both simulation cases lead to different time-dependent and spatial correlations among the two dimensions of the response, due to the use of different sets of Fourier basis in the generation of the multivariate eigenfunction in two dimensions.

A.5 Simulation Set-up for Multivariate Response in Three Dimensions

The independent multivariate response in three dimensions is generated as $X_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}^{(j)}\psi_\ell^{(j)}(t) + \epsilon_i^{(j)}(t)$ for $i = 1, \ldots, n$, $j = 1, 2, 3$. For generation of the univariate eigenfunctions, we consider 13 Fourier basis functions on the interval $\mathcal{T} = [0, 3]$. While $\psi_\ell^{(1)}(t)$ is taken to be the 1st, 6th and 7th basis functions evaluated on $[0, 1]$ (i.e., $L = 3$), $\psi_\ell^{(2)}(t)$ equals the 1st, 6th and 12th basis functions evaluated on $[1, 2]$, shifted to the left by one and multiplied by a random sign, and $\psi_\ell^{(3)}(t)$ equals the 6st, 7th and 12th basis functions evaluated on $[2, 3]$, shifted to the left by two and multiplied by a random sign. Multivariate PC scores, $\boldsymbol{\rho}_\ell^{(j)} = (\rho_{1\ell}^{(j)}, \rho_{2\ell}^{(j)}, \ldots, \rho_{n\ell}^{(j)})^\top$, along each dimension, are generated independently from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell^{(j)}(D - \nu W)^{-1}$, where the variance parameters $\alpha_\ell^{(j)}$'s are set to $\alpha_1^{(1)} = \alpha_1^{(2)} = \alpha_1^{(3)} = 2.0$, $\alpha_2^{(1)} = \alpha_2^{(2)} = \alpha_2^{(3)} = 1.2$,

$\alpha_3^{(1)} = \alpha_3^{(2)} = \alpha_3^{(3)} = 0.4$ and $\nu = 0.9$. The measurement errors, $\epsilon_i^{(j)}(t)$, are generated from $N(0, \sigma^2)$, where σ^2 is set to 0.02 and 0.5 corresponding low and high noise, respectively.

In the second more general correlated multivariate response in three dimensions the trivariate functional data is generated from the MST-FPCA model $X_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}\psi_\ell^{(j)}(t) + \epsilon_i^{(j)}(t)$ for $i = 1, \ldots, n$ and $j = 1, 2, 3$. The same basis sets as the first set-up are used. Distinct from the first set-up, the dependency of the multivariate response is induced by the shared PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\top$, generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$, where $\alpha_1 = 2.0$, $\alpha_2 = 1.2$, $\alpha_3 = 0.4$ and $\nu = 0.9$. The measurement error is simulated the same as above.

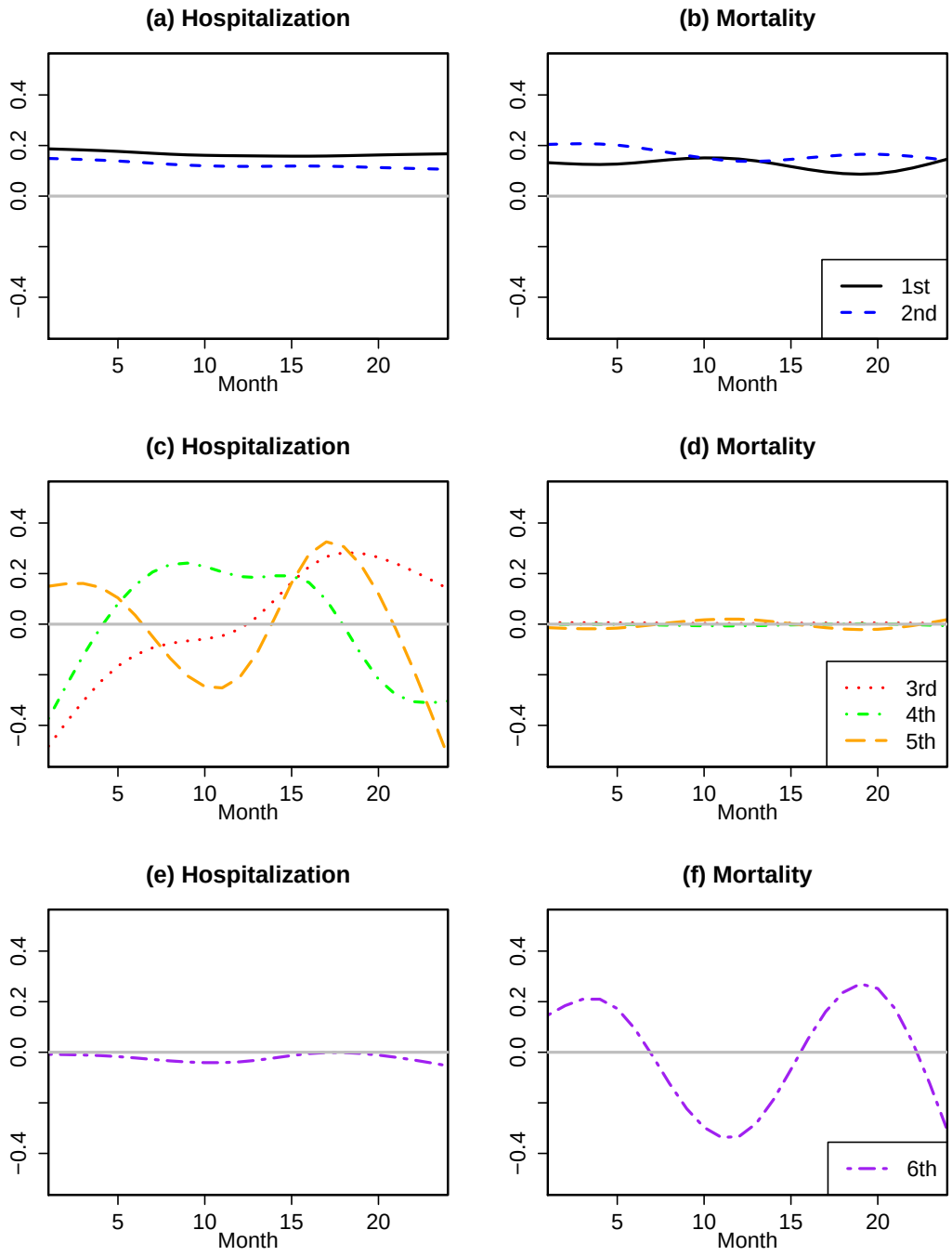

Figure A.1: Estimated multivariate eigenfunctions for hospitalization (a, c, e) and mortality (b, d, f). The two leading multivariate eigenfunctions describe overall constant variation in both hospitalization (a) and mortality (b) rates. The third, fourth and fifth multivariate eigenfunctions mainly explain variation in hospitalization (c), while the last multivariate eigenfunction mainly explains variation in mortality (f). The zero horizontal lines are included in gray for reference.

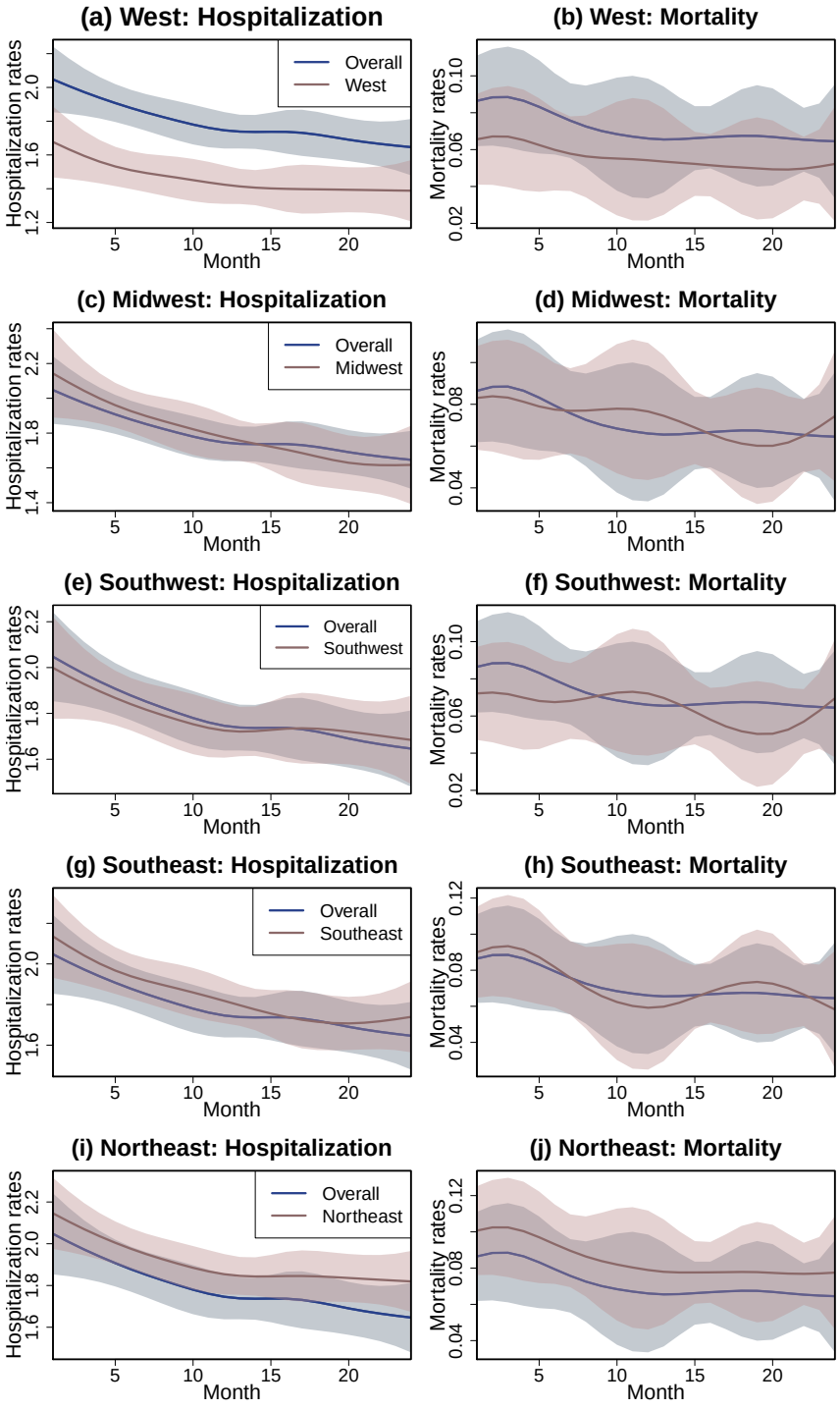

Figure A.2: Predicted hospitalization and mortality trajectories for five HSAs with median rates from the West (a-b), Midwest (c-d), Southwest (e-f), Southeast(g-h), and Northwest (i-j) zones (in red). Also provided are the 95% pointwise CIs (shaded) and predictions for a representative HSA with median rates across the entire U.S. (in blue) for comparison.

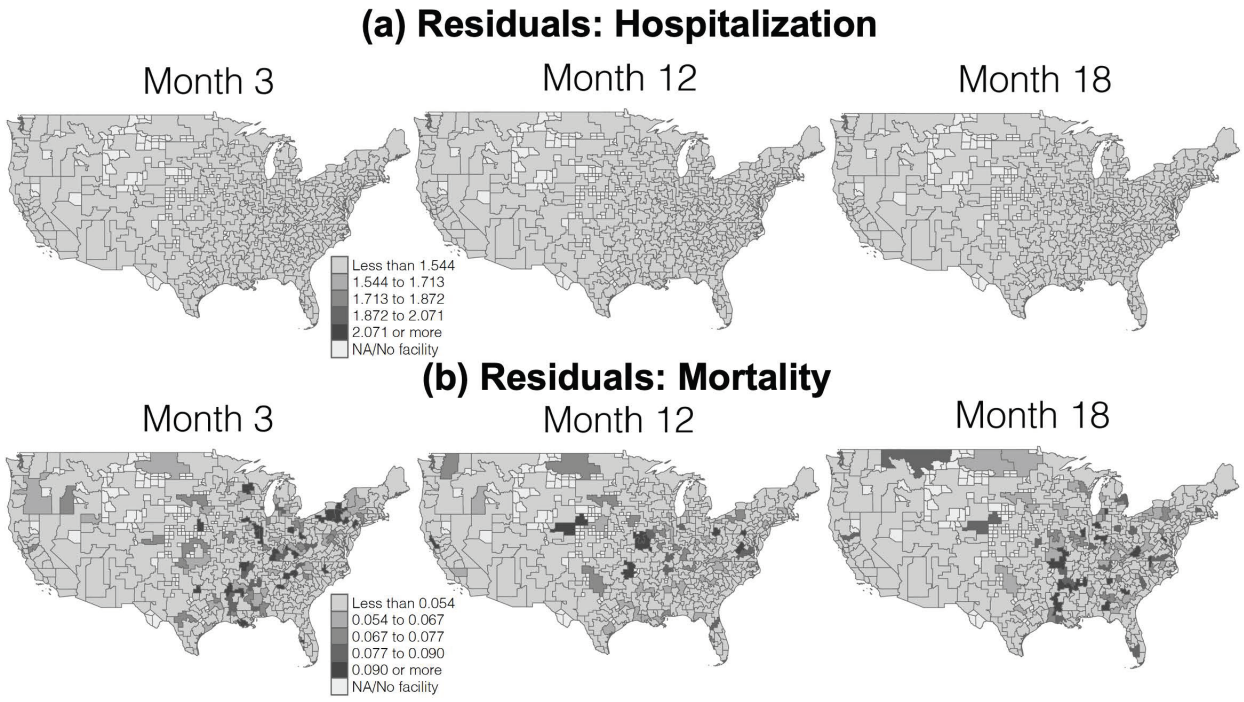

Figure A.3: The residuals (averaged across time) for hospitalization (a) and mortality (b) rates obtained from MST-FPCA fit to the data.

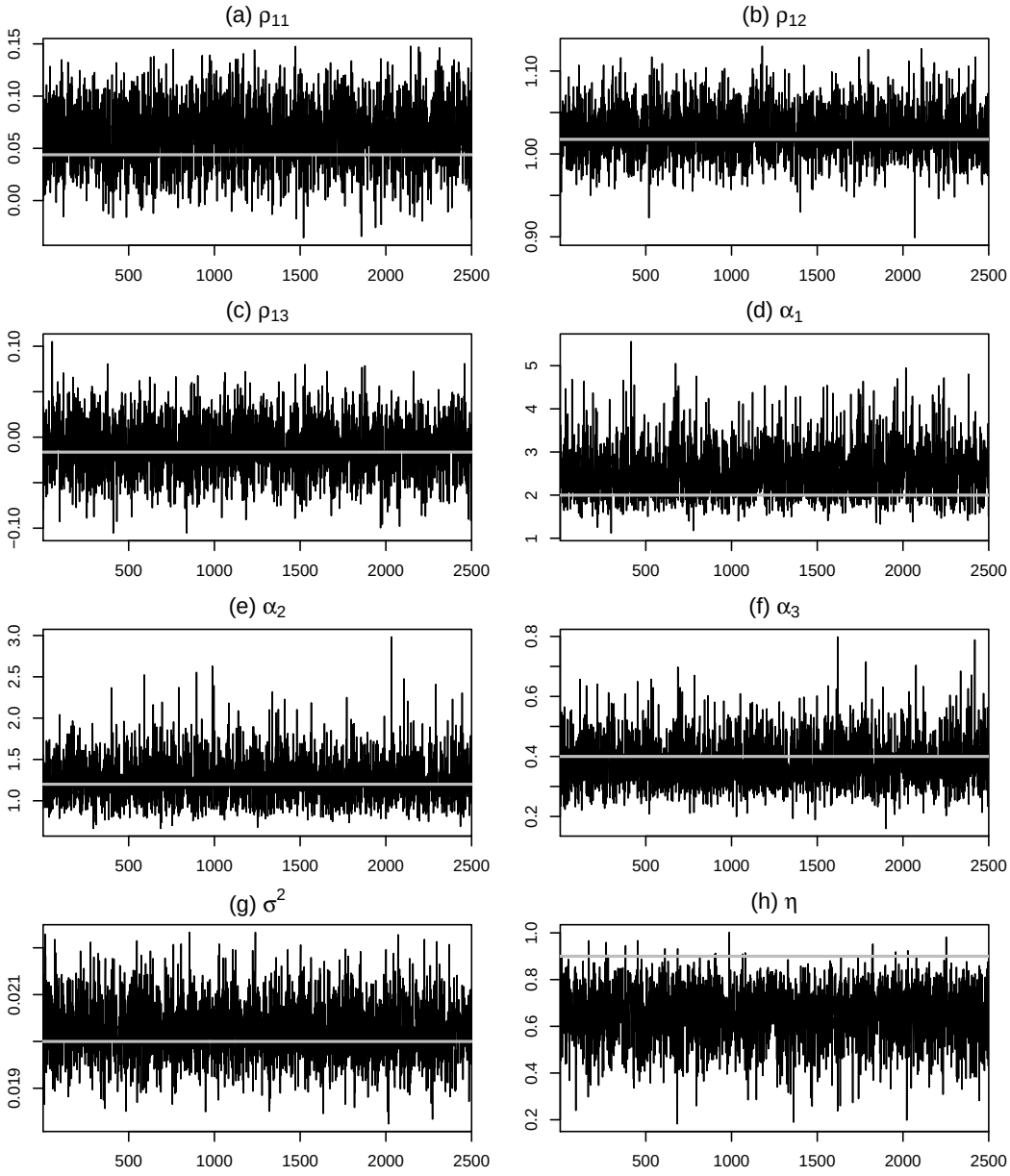

Figure A.4: MCMC trace plots for parameters in the second MCMC sampling (Step 4 in the estimation algorithm) in simulation studies on two-dimensional multivariate outcomes. For each chain, 3500 MCMC iterations are used where first 1000 iterations are discarded as burn-in. The true values of model parameters are represented by the gray horizontal lines.

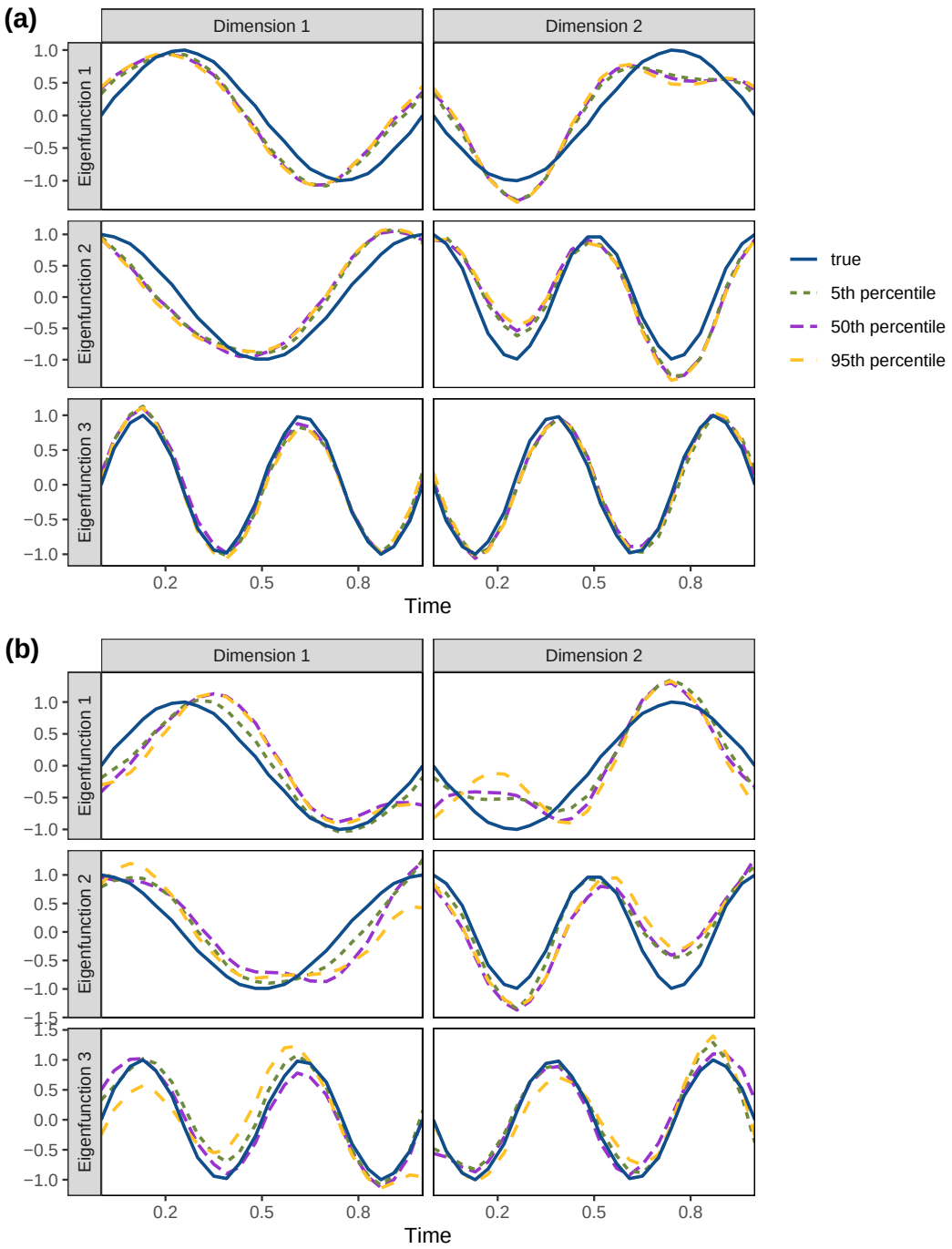

Figure A.5: Estimated eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs, from the correlated two-dimensional response set-up with number of time points $T = 24$, number of regions $n = 49$ and varying error variances $\sigma^2 = 0.02$ (a) and 0.5 (b) based on 200 Monte Carlo runs. The true eigenfunctions are given in solid blue.

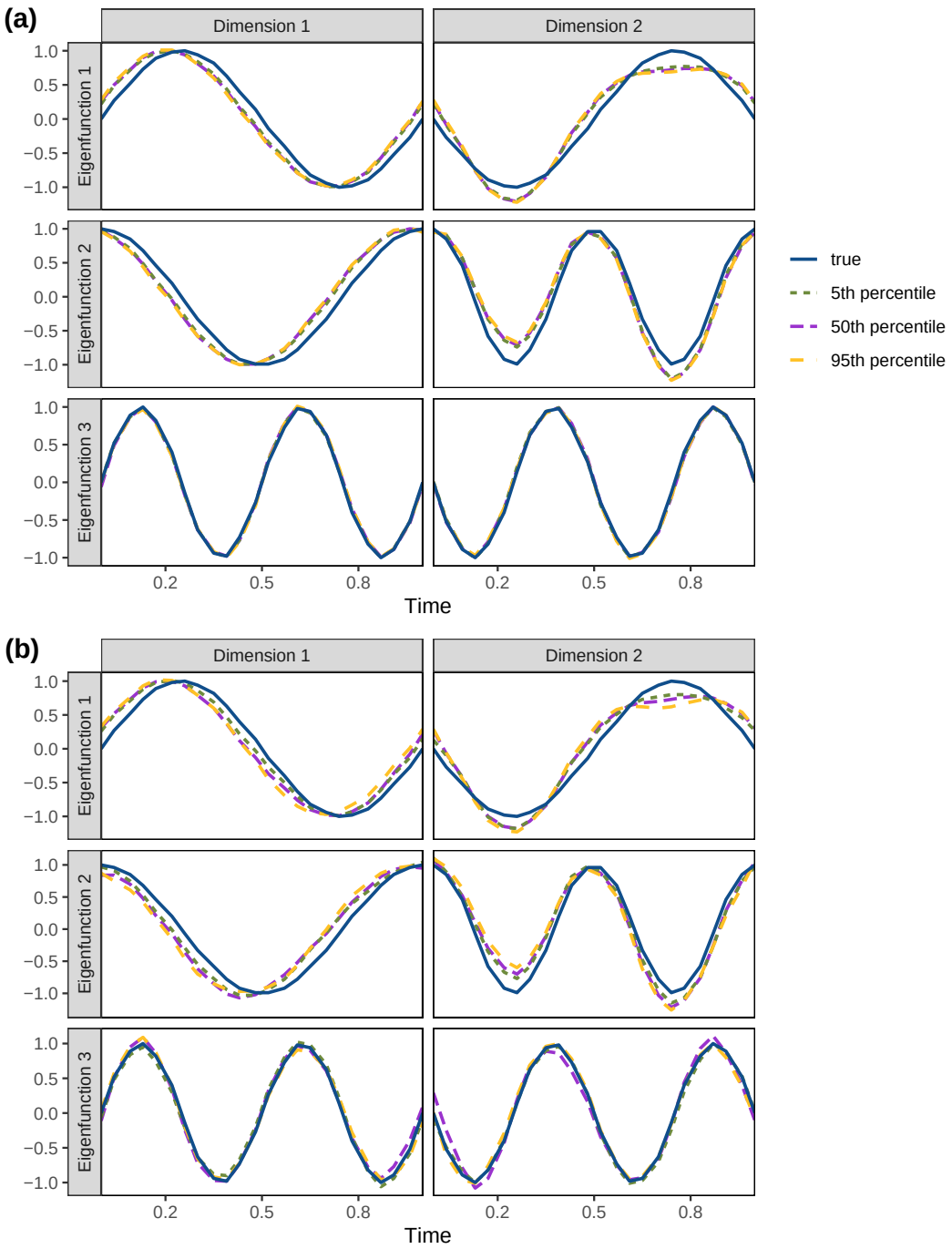

Figure A.6: Estimated eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs, from the correlated two-dimensional response set-up with number of time points $T = 24$, number of regions $n = 367$ and varying error variances $\sigma^2 = 0.02$ (a) and 0.5 (b) based on 200 Monte Carlo runs. The true eigenfunctions are given in solid blue.

Table A.1: The median and the 5th and 95th percentiles of the predicted region-specific hospitalization and mortality rates (with standard errors given in parentheses) from months 3, 12, and 18 on dialysis across the five zones: West, Midwest, Southwest, Southeast and Northeast. Also reported are overall rates across all health service areas (HSAs).

Month	**Percentile**	**Overall**	**West**	**Midwest**	**Southwest**	**Southeast**	**Northeast**
				Hospitalization			
	5	1.459(.069)	1.247(.071)	1.483(.077)	1.644(.079)	1.608(.072)	1.657(.091)
3	50	1.994(.071)	1.580(.082)	1.989(.069)	1.979(.073)	2.055(.074)	2.206(.075)
	95	2.592(.073)	1.997(.078)	2.638(.075)	2.367(.074)	2.652(.071)	2.539(.072)
	5	1.300(.060)	1.165(.056)	1.313(.061)	1.440(.055)	1.415(.055)	1.455(.057)
12	50	1.744(.053)	1.424(.058)	1.743(.057)	1.749(.054)	1.805(.063)	1.869(.056)
	95	2.240(.063)	1.713(.066)	2.290(.059)	2.039(.061)	2.294(.055)	2.212(.067)
	5	1.263(.082)	1.120(.074)	1.285(.078)	1.375(.075)	1.394(.071)	1.556(.087)
18	50	1.715(.071)	1.396(.073)	1.724(.073)	1.697(.087)	1.749(.078)	1.844(.073)
	95	2.184(.078)	1.655(.075)	2.208(.075)	1.953(.078)	2.190(.071)	2.165(.071)
				Mortality			
	5	.047(.013)	.041(.014)	.053(.014)	.047(.014)	.060(.013)	.065(.014)
3	50	.088(.014)	.067(.014)	.088(.014)	.082(.014)	.089(.014)	.101(.014)
	95	.134(.014)	.100(.013)	.133(.014)	.115(.014)	.139(.014)	.139(.014)
	5	.035(.016)	.034(.017)	.035(.016)	.031(.016)	.038(.016)	.057(.017)
12	50	.068(.017)	.057(.017)	.066(.016)	.062(.017)	.069(.017)	.083(.016)
	95	.107(.016)	.079(.015)	.100(.017)	.096(.016)	.107(.016)	.114(.017)
	5	.033(.013)	.027(.013)	.035(.013)	.035(.012)	.042(.013)	.048(.013)
18	50	.066(.012)	.050(.013)	.067(.012)	.059(.013)	.067(.013)	.075(.013)
	95	.106(.013)	.077(.013)	.107(.013)	.090(.013)	.109(.012)	.108(.013)

Table A.2: The mean MSE and MSDE values for estimation of the model components from the correlated three-dimensional response simulation set-up with number of time points $T = 24$, and varying number of regions n and measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	**49 regions**		**367 regions**	
Noise level, σ^2	0.02	0.5	0.02	0.5
	MSDE			
$\widehat{\psi}_1^{(1)}(t)$	0.061	0.077	0.016	0.018
$\widehat{\psi}_1^{(2)}(t)$	0.062	0.075	0.016	0.017
$\widehat{\psi}_1^{(3)}(t)$	0.060	0.074	0.016	0.018
$\widehat{\psi}_2^{(1)}(t)$	0.056	0.075	0.016	0.018
$\widehat{\psi}_2^{(2)}(t)$	0.056	0.079	0.016	0.018
$\widehat{\psi}_2^{(3)}(t)$	0.057	0.076	0.016	0.019
$\widehat{\psi}_3^{(1)}(t)$	0.012	0.034	<0.001	0.004
$\widehat{\psi}_3^{(2)}(t)$	0.011	0.043	0.001	0.005
$\widehat{\psi}_3^{(3)}(t)$	0.011	0.046	0.001	0.005
	MSE			
$\widehat{\alpha}_1$	0.059	0.064	0.056	0.061
$\widehat{\alpha}_2$	0.038	0.040	0.013	0.020
$\widehat{\alpha}_3$	0.099	0.143	0.045	0.033
$\widehat{\nu}$	0.284	0.287	0.153	0.049
$\widehat{\sigma}^2$	< 0.001	0.002	< 0.001	< 0.001
$\widehat{\rho}_{i1}$	0.470	0.493	0.039	0.068
$\widehat{\rho}_{i2}$	0.287	0.383	0.101	0.151
$\widehat{\rho}_{i3}$	0.157	0.316	0.010	0.139

Table A.3: The mean MSDE of the predicted multivariate region-specific trajectories and the coverage probability (CP) and length of the associated 95% pointwise credible intervals (CIs) based on M-FPCA, MBYM, U-FPCA and MST-FPCA from both independent and correlated three-dimensional response simulation set-ups with number of time points $T = 24$, number of regions $n = 367$, and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	367 **regions**							
Noise level, σ^2	0.02				0.5			
Model	M-FPCA	MBYM	U-FPCA	MST-FPCA	M-FPCA	MBYM	U-FPCA	MST-FPCA
	Independent response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.094	0.406	0.023	0.023	0.489	0.547	0.376	0.387
MSDE: $\widehat{x}_i^{(2)}(t)$	0.086	0.506	0.019	0.019	0.497	0.674	0.326	0.351
MSDE: $\widehat{x}_i^{(3)}(t)$	0.102	0.901	0.012	0.012	0.561	1.063	0.201	0.219
Length$^{(1)}$	0.195	0.341	0.194	0.195	0.793	0.602	0.841	0.853
Length$^{(2)}$	0.190	0.332	0.191	0.191	0.790	0.607	0.831	0.842
Length$^{(3)}$	0.192	0.367	0.193	0.193	0.792	0.492	0.848	0.842
CP$^{(1)}$ (%)	88.02	37.48	94.82	94.81	88.30	56.34	94.77	94.35
CP$^{(2)}$ (%)	87.02	44.88	93.75	93.69	87.85	62.45	94.48	94.17
CP$^{(3)}$ (%)	87.78	24.14	94.14	94.04	88.03	31.75	94.56	94.26
	Correlated response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.096	0.405	0.023	0.015	0.432	0.520	0.385	0.206
MSDE: $\widehat{x}_i^{(2)}(t)$	0.097	0.406	0.023	0.015	0.428	0.560	0.382	0.206
MSDE: $\widehat{x}_i^{(3)}(t)$	0.098	1.033	0.023	0.015	0.421	1.175	0.364	0.210
Length$^{(1)}$	0.165	0.318	0.194	0.174	0.664	0.550	0.849	0.696
Length$^{(2)}$	0.166	0.316	0.190	0.170	0.665	0.550	0.832	0.677
Length$^{(3)}$	0.166	0.360	0.193	0.172	0.665	0.488	0.844	0.690
CP$^{(1)}$ (%)	91.61	32.42	94.70	96.25	92.47	50.70	94.64	97.20
CP$^{(2)}$ (%)	91.60	41.38	94.28	95.78	92.46	56.50	94.52	96.96
CP$^{(3)}$ (%)	91.56	23.21	94.33	95.80	92.43	31.16	94.51	97.02
	Run times							
	M-FPCA		MBYM		U-FPCA		MST-FPCA	
Time(min)	10.603		145.503		11.843		24.338	

Table A.4: The mean MSDE of the predicted multivariate region-specific trajectories and the coverage probability (CP) and length of the associated 95% pointwise credible intervals (CIs) based on M-FPCA, MBYM, U-FPCA and MST-FPCA from both independent and correlated two-dimensional response simulation set-ups with number of time points $T = 24$, number of regions $n = 49$, and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	49 **regions**							
Noise level, σ^2	0.02				0.5			
Model	M-FPCA	MBYM	U-FPCA	MST-FPCA	M-FPCA	MBYM	U-FPCA	MST-FPCA
	Independent response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.073	1.364	0.026	0.026	0.532	1.439	0.421	0.431
MSDE: $\widehat{x}_i^{(2)}(t)$	0.079	0.519	0.021	0.023	0.516	0.658	0.367	0.376
Length$^{(1)}$	0.193	0.575	0.196	0.193	0.839	0.717	0.870	0.875
Length$^{(2)}$	0.191	0.423	0.191	0.192	0.835	0.681	0.857	0.858
CP$^{(1)}$ (%)	86.39	29.49	89.94	89.79	86.97	36.73	92.12	92.21
CP$^{(2)}$ (%)	86.88	44.04	91.04	91.24	86.90	56.12	91.56	91.78
	Correlated response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.086	1.364	0.027	0.024	0.515	1.460	0.287	0.235
MSDE: $\widehat{x}_i^{(2)}(t)$	0.086	1.376	0.027	0.024	0.516	1.468	0.300	0.248
Length$^{(1)}$	0.158	0.574	0.195	0.191	0.739	0.692	0.860	0.818
Length$^{(2)}$	0.158	0.572	0.194	0.190	0.738	0.689	0.856	0.817
CP$^{(1)}$ (%)	88.96	29.51	90.03	90.22	89.03	40.37	92.04	93.46
CP$^{(2)}$ (%)	88.00	31.53	89.53	89.77	88.83	42.78	91.15	92.78
	Run times							
	M-FPCA		MBYM		U-FPCA		MST-FPCA	
Time(min)	0.603		8.079		0.632		1.314	

Table A.5: The mean MSDE of the predicted multivariate region-specific trajectories and the coverage probability (CP) and length of the associated 95% pointwise credible intervals (CIs) based on M-FPCA, MBYM, U-FPCA and MST-FPCA from both independent and correlated three-dimensional response simulation set-ups with number of time points $T = 24$, number of regions $n = 49$ and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	49 **regions**							
Noise level, σ^2	0.02				0.5			
Model	M-FPCA	MBYM	U-FPCA	MST-FPCA	M-FPCA	MBYM	U-FPCA	MST-FPCA
	Independent response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.111	0.555	0.030	0.030	0.635	0.738	0.491	0.527
MSDE: $\widehat{x}_i^{(2)}(t)$	0.119	0.572	0.027	0.027	0.650	0.697	0.446	0.475
MSDE: $\widehat{x}_i^{(3)}(t)$	0.118	1.009	0.020	0.020	0.667	1.036	0.340	0.367
Length$^{(1)}$	0.193	0.357	0.194	0.194	0.774	0.633	0.826	0.837
Length$^{(2)}$	0.188	0.349	0.189	0.190	0.755	0.632	0.811	0.826
Length$^{(3)}$	0.191	0.502	0.194	0.192	0.749	0.669	0.830	0.846
CP$^{(1)}$ (%)	86.38	38.87	92.27	92.36	86.97	57.48	92.06	92.37
CP$^{(2)}$ (%)	85.60	43.67	91.47	91.64	85.72	60.28	91.21	91.62
CP$^{(3)}$ (%)	86.00	32.21	89.84	89.61	86.32	41.72	91.44	91.46
	Correlated response							
MSDE: $\widehat{x}_i^{(1)}(t)$	0.105	0.548	0.023	0.021	0.502	0.649	0.385	0.299
MSDE: $\widehat{x}_i^{(2)}(t)$	0.105	0.547	0.023	0.020	0.503	0.646	0.405	0.352
MSDE: $\widehat{x}_i^{(3)}(t)$	0.105	1.417	0.024	0.022	0.503	1.449	0.391	0.313
Length$^{(1)}$	0.154	0.352	0.193	0.188	0.635	0.607	0.811	0.787
Length$^{(2)}$	0.154	0.349	0.189	0.184	0.635	0.621	0.814	0.771
Length$^{(3)}$	0.154	0.458	0.192	0.187	0.636	0.644	0.822	0.786
CP$^{(1)}$ (%)	87.83	38.28	92.27	92.77	88.98	56.85	91.05	94.88
CP$^{(2)}$ (%)	88.21	47.35	90.84	91.29	88.90	62.21	91.10	93.57
CP$^{(3)}$ (%)	87.93	34.28	90.44	90.58	89.07	44.01	90.44	93.49
	Run times							
	M-FPCA		MBYM		U-FPCA		MST-FPCA	
Time(min)	0.912		12.306		0.994		2.016	

Appendix B

Supporting Information for Chapter 2

B.1 Details on the USRDS Study Cohort

The main eligibility criteria for the study cohort are that (a) patients survive the first 90 days of dialysis and do not recover renal function or receive kidney transplantation and (b) that they have Medicare as their primary payer on day 91 of dialysis. To obtain stable region-specific estimation and inference, HSAs are merged first to guarantee that they contain at least 5 facilities and for a second time to guarantee that the total number of months with zero mortality rate does not exceed one-third of the two year follow-up. Both merging algorithms have four steps: Step 1) start with HSAs as regions; Step 2) among regions with the lowest number of facilities (or with greater than eight months of zero mortality), randomly select one region (call it R); Step 3) among neighbors of the selected region R, randomly select a neighbor, among those with the lowest number of facilities (or among those with the largest number of months with zero mortality rate), to merge with region R; Step 4) repeat Steps 2 and 3 until the lowest number of facilities within a region is 5 (or the largest number of months with zero mortality rate during the two-year follow up within a region is eight).

B.2 Simulation Set-up for Multivariate Response in Three Dimensions

In the simulation setting with three-dimensional multivariate response, the trivariate response is simulated using

$$Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}^{(j)}(t_k) + \sum_{\ell=1}^{L} \rho_{i\ell}\psi_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k), \text{ for } j = 1, 2, 3,$$

where t_k, $k = 1, \ldots, 24$, denote the equidistant grid of time points between 0 and 1, mimicking the 24 month follow-up in our data application. The covariates $\boldsymbol{X}_i = \{1, X_{i1}, X_{i2}\}^{\mathsf{T}}$ for $i = 1, \ldots, n$ include a y-intercept term, and X_{i1}, X_{i2} generated from a bivariate normal with mean $(0, 0)^{\mathsf{T}}$, variance 1 and covariance $2^{-0.5}$. The time-varying coefficient function, $\boldsymbol{\beta}^{(j)}(t_k) = \{\beta_1^{(j)}(t_k), \beta_2^{(j)}(t_k), \beta_3^{(j)}(t_k)\}$, $j = 1, 2, 3$, is set as follows:

$$\beta_0^{(1)}(t) = 0.1\exp(2t-1), \quad \beta_1^{(1)}(t) = 0.25t(1-t), \quad \beta_2^{(1)}(t) = 0.06(1-t)^2,$$

$$\beta_0^{(2)}(t) = 0.15\sqrt{2t} + 0.05, \quad \beta_1^{(2)}(t) = 0.25(t-0.5)^2, \quad \beta_2^{(2)}(t) = 0.06t^2,$$

$$\beta_0^{(3)}(t) = 0.15t + 0.05, \quad \beta_1^{(3)}(t) = 0.06t^2, \quad \beta_2^{(3)}(t) = 0.05\sqrt{0.5t}.$$

The multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t_k) = \{\psi_\ell^{(1)}(t_k), \psi_\ell^{(2)}(t_k), \psi_\ell^{(3)}(t_k)\}^{\mathsf{T}}$, $\ell = 1, 2, 3$ (for $L = 3$), are generated using 13 Fourier basis functions on the interval $\mathcal{T} = [0, 3]$. While $\psi_\ell^{(1)}(t)$ is taken to be the 12th, 6th and 7th basis functions evaluated on $[0, 1]$ (i.e., $L = 3$), $\psi_\ell^{(2)}(t)$ equals the 7th, 6th and 12th basis functions evaluated on $[1, 2]$, shifted to the left by one and multiplied by a random sign, and $\psi_\ell^{(3)}(t)$ equals the 6th, 12th and 7th basis functions evaluated on $[2, 3]$, shifted to the left by two and multiplied by a random sign. Similar to the simulation with the two-dimensional multivariate response, the multivariate PC scores, $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\top$ are generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$, where W is the adjacency matrix and D is the diagonal matrix utilizing the map of $n = 49$ states in the contiguous U.S. (including the District of Columbia). The variance parameters α_ℓ's are taken to be $\alpha_1 = 2.4$, $\alpha_2 = 1.2$, $\alpha_3 = 0.6$, while the spatial correlation parameter is set to $\nu = 0.97$. Finally, the measurement errors, $\boldsymbol{\epsilon_i}(t_k) = \{\epsilon_i^{(1)}(t_k), \epsilon_i^{(2)}(t_k), \epsilon_i^{(3)}(t_k)\}^{\mathsf{T}} \sim N\left((0, 0, 0)^{\mathsf{T}}, \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)\right)$, where $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ equal 0.2 and 2 corresponding to low and high noise simulation cases, respectively.

B.3 Priors Used in the Data Analysis and Simulation

In Step 3 of the proposed MV-VCSTM estimation algorithm, the between-score covariance matrix Σ_b is targeted through the modeling framework:

$$\widehat{\boldsymbol{\xi}} = \boldsymbol{\xi} + \mathbf{e}$$

$$\mathbf{e} \sim N(\mathbf{0}, \tau^2 I), \;\; \boldsymbol{\xi} \sim N\left(\mathbf{0}, AA^{\mathsf{T}} \otimes (D - \nu W)^{-1}\right),$$

$$a_{\ell\ell} \sim \text{Lognormal}(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}}), \;\; a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}}), \;\; \nu \sim \text{Unif}(a_\nu, b_\nu), \;\; \tau^2 \sim \text{IG}(a_{\tau^2}, b_{\tau^2}),$$

where elements of $\widehat{\boldsymbol{\xi}}$ are the scores estimated from univariate FPCA. Elementwise priors are imposed on A where $a_{\ell\ell} \sim \text{lognormal}(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}})$, and $a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$ for $1 \leq \ell' < \ell$, with $\mu_{a_{\ell\ell'}} = -3$, $\sigma^2_{a_{\ell\ell'}} = 1$, $\mu_{a_{\ell\ell'}} = 0$, and $\sigma^2_{a_{\ell\ell'}} = 5$. We apply an Inverse Gamma (IG)(a_{τ^2}, b_{τ^2}) prior for τ^2 ($a_{\tau^2} = b_{\tau^2} = 1$ in our application), and a Uniform prior for the spatial parameter ν with parameters constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues (denoted by a_ν, b_ν, respectively) of the matrix $D^{-1/2}WD^{-1/2}$.

In Step 5 of the MV-VCSTM estimation algorithm, we target the varying coefficient functions (VCFs) $\boldsymbol{\beta}^{(j)}(t)$, multivariate PC scores ($\rho_{i\ell}$), spatial parameters (α_ℓ and ν), and measurement error variance (σ^2) of MV-VCSTM via expanding the VCFs on thin-plate splines and the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions $\widehat{\boldsymbol{\psi}}_\ell(t)$ as follows:

$$Y_i^{(j)}(t_k) = \sum_{p=1}^{P} X_{ip} \left(u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq} \right) + \sum_{\ell=1}^{M^+} \rho_{i\ell} \widehat{\psi}_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k) \quad \text{for} \;\; j = 1, \ldots, J,$$

$$\epsilon_i^{(j)}(t_k) \sim N(0, \sigma_j^2), \quad \boldsymbol{\rho}_\ell \sim N(\mathbf{0}, \alpha_\ell (D - \nu W)^{-1}), \quad \alpha_\ell \sim \text{IG}(a_{\alpha_\ell}, b_{\alpha_\ell}), \quad \nu \sim \text{Unif}(a_\nu, b_\nu),$$

$$u_{p0}^{(j)} \sim N(0, \sigma^2_{u_{p0j}}), \quad u_{p1}^{(j)} \sim N(0, \sigma^2_{u_{p1j}}), \quad \widetilde{v}_{pq}^{(j)} \sim N(0, \sigma^2_{\widetilde{v}_{pj}}),$$

$$\sigma^2_{\widetilde{v}_{pj}} \sim \text{IG}(a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}}), \quad \sigma_j^2 \sim \text{IG}(a_{\sigma_j^2}, b_{\sigma_j^2}),$$

where $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^{\mathsf{T}}$ denotes the $n \times 1$ score vector for ℓth eigencomponent. Non-

informative or weak informative priors are used for $u_{p0}^{(j)}$, $u_{p1}^{(j)}$, $\widetilde{v}_{pq}^{(j)}$, $\epsilon_i^{(j)}(t)$ and α_ℓ, where $\sigma^2_{u_{p0j}} = \sigma^2_{u_{p1j}} = 10^6$, $a_{\sigma^2_{\widetilde{v}_{pj}}} = b_{\sigma^2_{\widetilde{v}_{pj}}} = 10^{-6}$, and $a_{\sigma_j^2} = b_{\sigma_j^2} = a_{\alpha_\ell} = b_{\alpha_\ell} = 1$. Finally, the prior imposed on ν is the same as in Step 3.

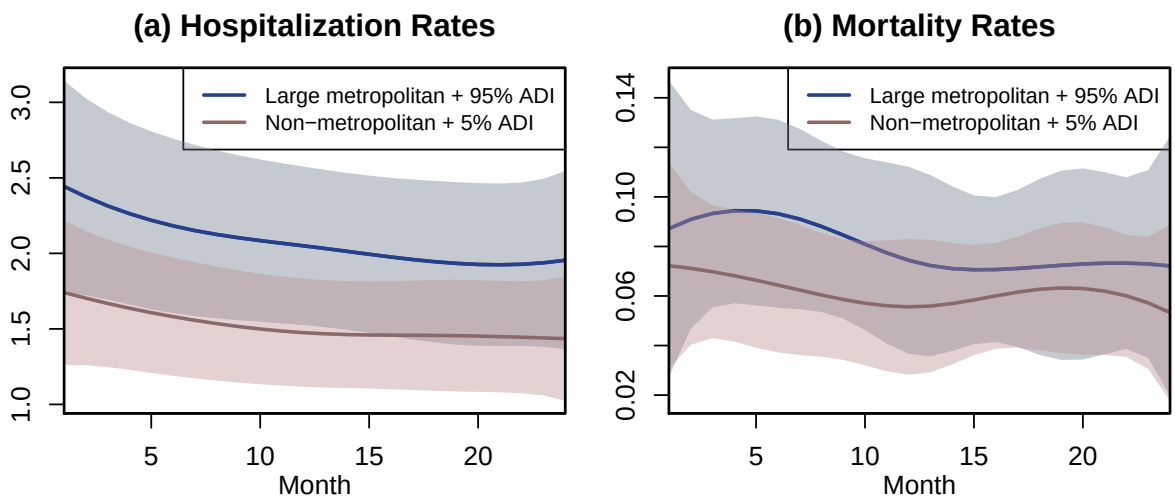

Figure B.1: Predicted hospitalization (a) and mortality (b) trajectories for two hypothetical regions, one of large metropolitan with 95% percentile ADI (blue) and the other one of non-metropolitan with 5% percentile ADI (red) while other covariates are kept at mean or reference levels, along with 95% simultaneous credible bands (shaded).

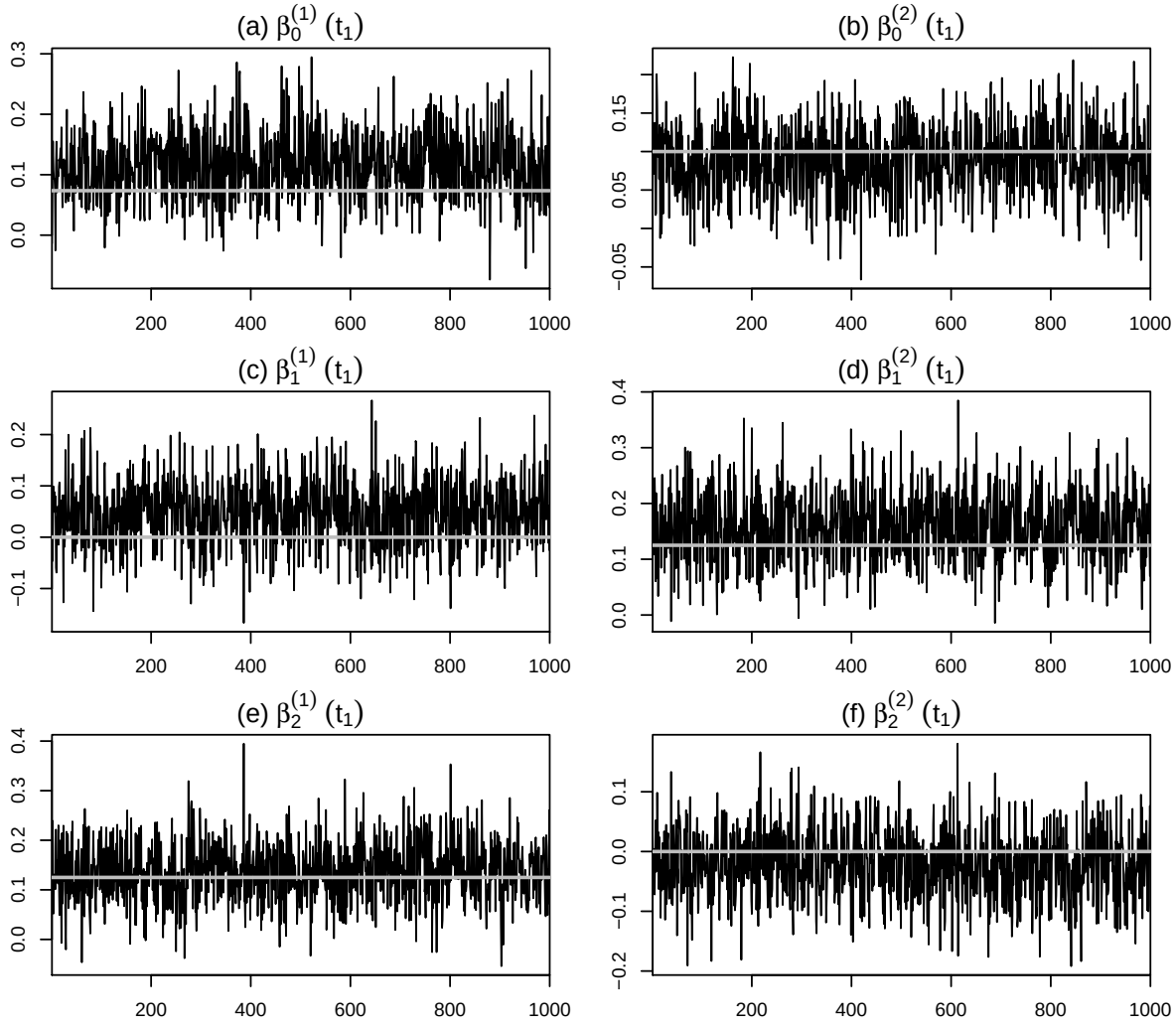

Figure B.2: MCMC trace plots for VCFs from Step 5 of the estimation algorithm, where 15000 iterations (with 5000 as burn-in) are used and one out of every 10 samples is returned. The true values of VCFs are represented by the gray horizontal lines. The results are based on two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variance $\sigma^2 = 0.2$.

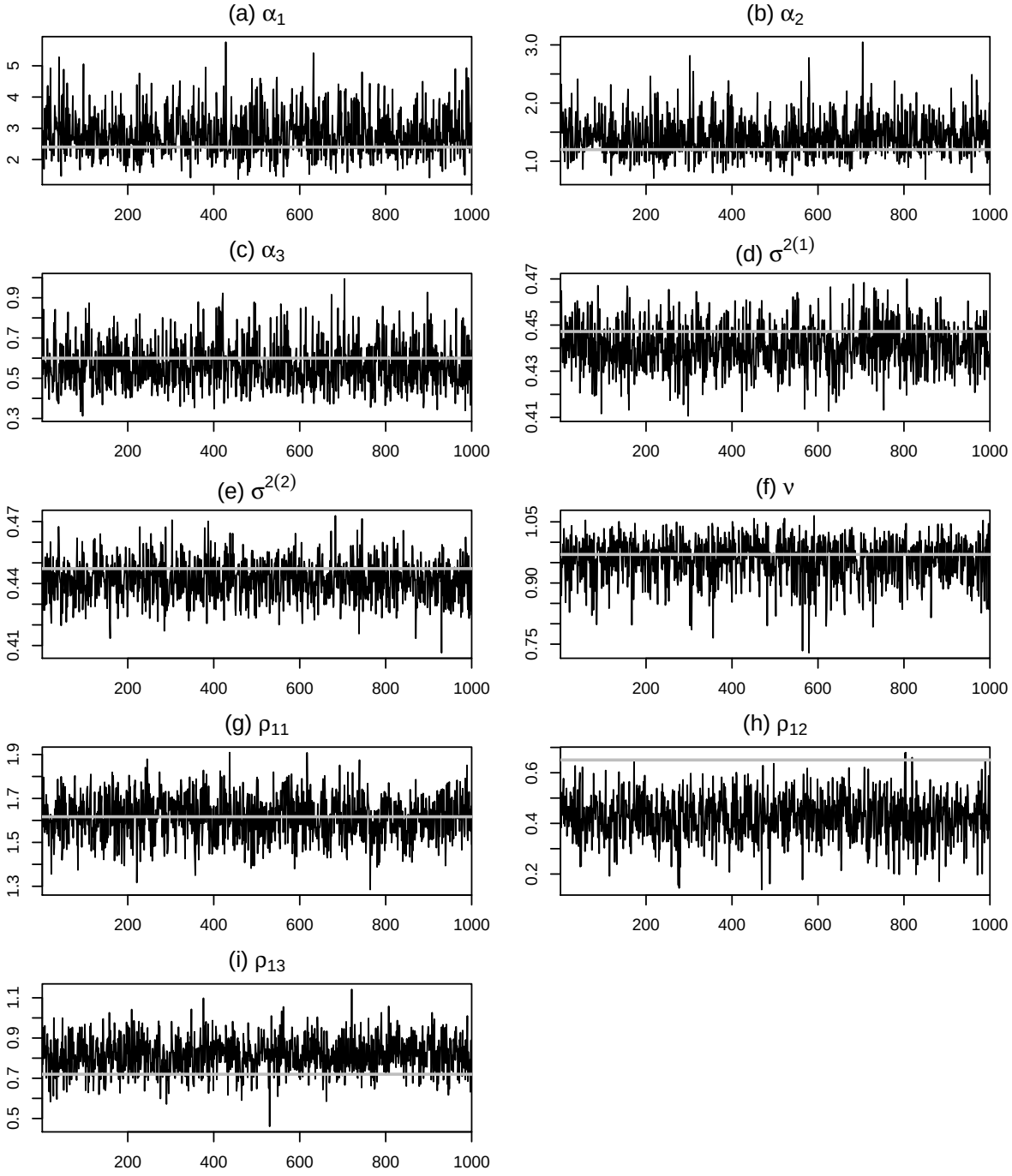

Figure B.3: MCMC trace plots for model parameters from Step 5 of the estimation algorithm, where 15000 iterations (with 5000 as burn-in) are used and one out of every 10 samples is returned. The true values of model parameters are represented by the gray horizontal lines. The results are based on two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variance $\sigma^2 = 0.2$.

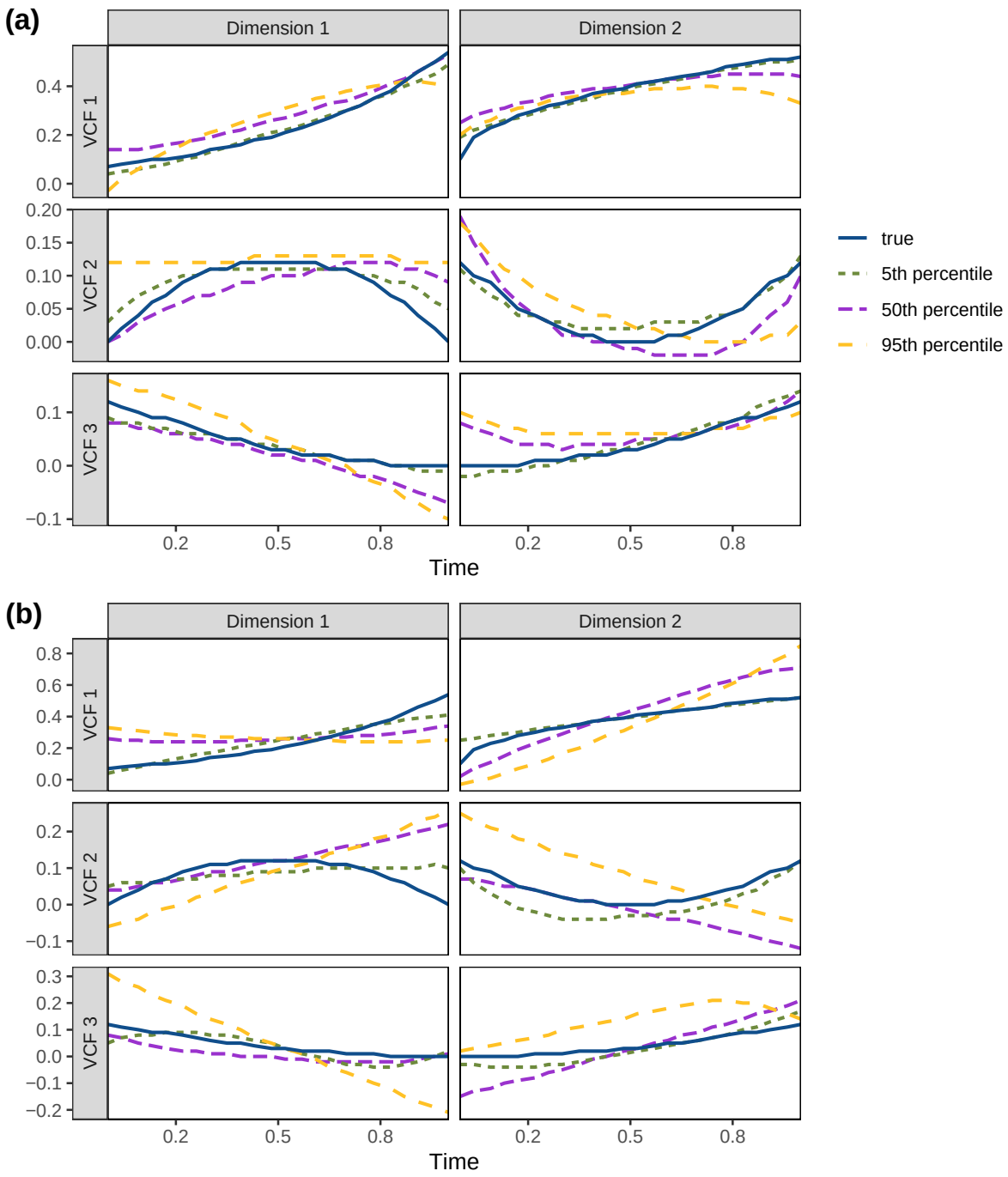

Figure B.4: Estimated VCFs from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the simulation set-up with two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variances $\sigma^2 = 0.2$ (a) and $\sigma^2 = 2$ (b). The true VCFs are given in solid blue.

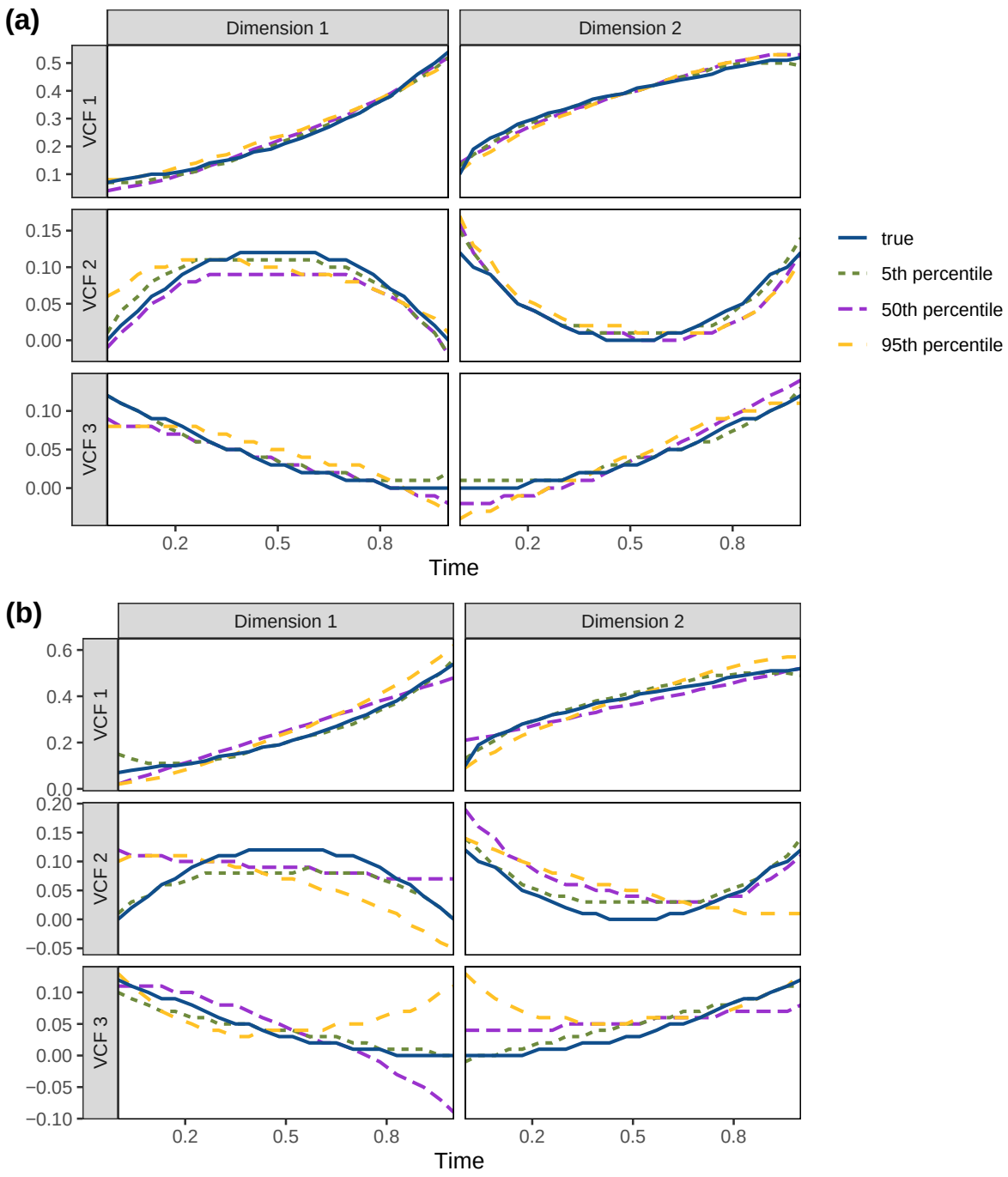

Figure B.5: Estimated VCFs from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the simulation set-up with two-dimensional multivariate response with aligned time points, $n = 367$ regions, and measurement error variances $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 2$ (b). The true VCFs are given in solid blue.

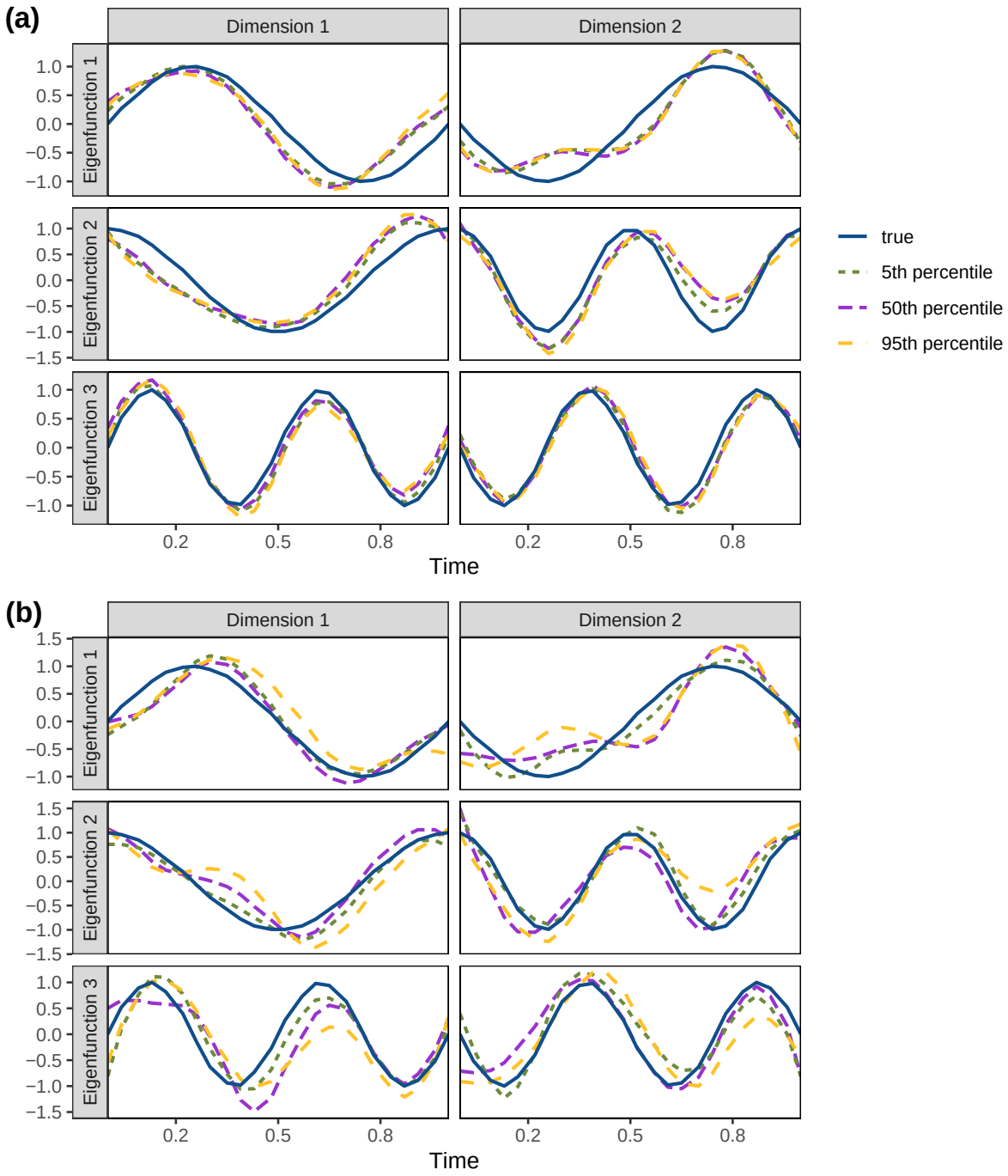

Figure B.6: Estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the set-up with two-dimensional multivariate response with aligned time points, $n = 49$ regions, and $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 2$ (b). The true multivariate eigenfunctions are given in solid blue.

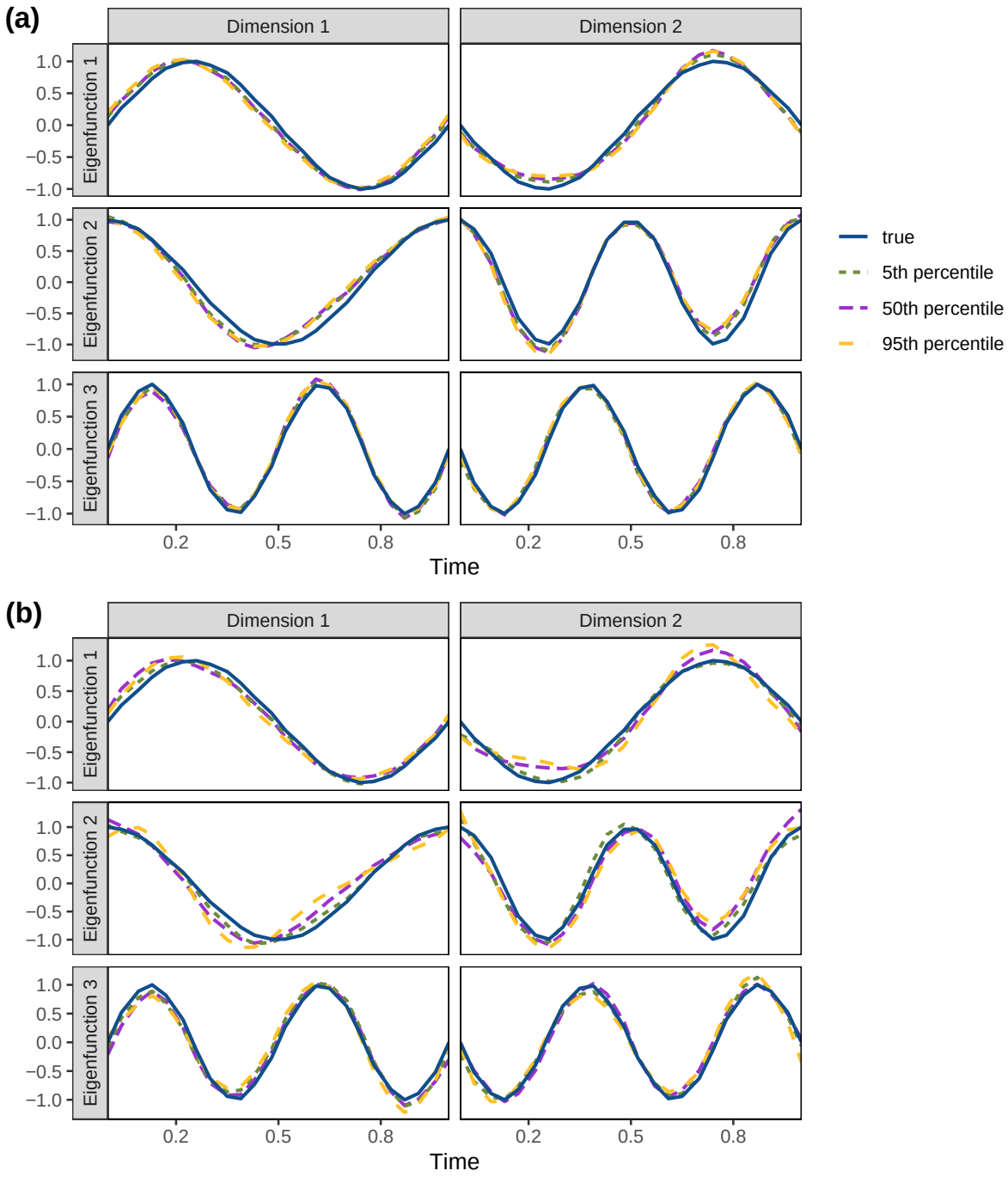

Figure B.7: Estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the set-up with two-dimensional multivariate response with aligned time points, $n = 367$ regions, and $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 0.2$ (b). The true multivariate eigenfunctions are given in solid blue.

Table B.1: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from four simulation settings with two sets of time points (aligned and misaligned) and two sets of measurement error variance σ^2. Results are based on two-dimensional multivariate response, $n = 49$ regions and 200 Monte Carlo runs.

Time points	**Aligned**		**Misaligned**	
Noise level, σ^2	0.2	2	0.2	2
	MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.0033	0.0178	0.0130	0.0590
$\widehat{\beta}_0^{(2)}(t)$	0.0031	0.0236	0.0107	0.0551
$\widehat{\beta}_1^{(1)}(t)$	0.0017	0.0096	0.0038	0.0095
$\widehat{\beta}_1^{(2)}(t)$	0.0013	0.0093	0.0036	0.0096
$\widehat{\beta}_2^{(1)}(t)$	0.0011	0.0126	0.0038	0.0089
$\widehat{\beta}_2^{(2)}(t)$	0.0009	0.0093	0.0034	0.0079
$\widehat{\psi}_1^{(1)}(t)$	0.0795	0.0872	0.0373	0.0513
$\widehat{\psi}_1^{(2)}(t)$	0.0821	0.0846	0.0330	0.0507
$\widehat{\psi}_2^{(1)}(t)$	0.0960	0.1222	0.0292	0.1013
$\widehat{\psi}_2^{(2)}(t)$	0.1060	0.1208	0.0272	0.1099
$\widehat{\psi}_3^{(1)}(t)$	0.0349	0.1697	0.0245	0.3649
$\widehat{\psi}_3^{(2)}(t)$	0.0333	0.1769	0.0283	0.3602
$\widehat{U}_i^{(1)}(t)$	0.0465	0.3447	0.2234	0.3773
$\widehat{U}_i^{(2)}(t)$	0.0470	0.3516	0.2207	0.3623
$\widehat{Y}_i^{(1)}(t)$	0.0251	0.3499	0.0398	0.2693
$\widehat{Y}_i^{(2)}(t)$	0.0205	0.2152	0.0278	0.2241
	MSE			
$\widehat{\alpha}_1$	0.0985	0.1067	0.0511	0.0993
$\widehat{\alpha}_2$	0.0260	0.0347	0.0168	0.0602
$\widehat{\alpha}_3$	0.0188	0.0696	0.0078	0.1126
$\widehat{\nu}$	0.0067	0.0070	0.0163	0.0836
$\widehat{\sigma}^{2(1)}$	0.0022	0.0023	0.0032	0.0034
$\widehat{\sigma}^{2(2)}$	0.0023	0.0024	0.0019	0.0021

Table B.2: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from three-dimensional multivariate response with aligned time points, $n = 49$ regions and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	**49 regions**	
Noise level, σ^2	0.2	2
	MSDE	
$\widehat{\beta}_0^{(1)}(t)$	0.0047	0.0576
$\widehat{\beta}_0^{(2)}(t)$	0.0059	0.0651
$\widehat{\beta}_0^{(3)}(t)$	0.0097	0.0582
$\widehat{\beta}_1^{(1)}(t)$	0.0012	0.0097
$\widehat{\beta}_1^{(2)}(t)$	0.0011	0.0090
$\widehat{\beta}_1^{(3)}(t)$	0.0011	0.0114
$\widehat{\beta}_2^{(1)}(t)$	0.0009	0.0092
$\widehat{\beta}_2^{(2)}(t)$	0.0008	0.0093
$\widehat{\beta}_2^{(3)}(t)$	0.0010	0.0098
$\widehat{\psi}_1^{(1)}(t)$	0.0600	0.0752
$\widehat{\psi}_1^{(2)}(t)$	0.0612	0.0689
$\widehat{\psi}_1^{(3)}(t)$	0.0616	0.0948
$\widehat{\psi}_2^{(1)}(t)$	0.0969	0.1401
$\widehat{\psi}_2^{(2)}(t)$	0.0929	0.1335
$\widehat{\psi}_2^{(3)}(t)$	0.1006	0.2084
$\widehat{\psi}_3^{(1)}(t)$	0.0649	0.2304
$\widehat{\psi}_3^{(2)}(t)$	0.0599	0.2700
$\widehat{\psi}_3^{(3)}(t)$	0.0559	0.1585
$\widehat{U}_i^{(1)}(t)$	0.0309	0.2417
$\widehat{U}_i^{(2)}(t)$	0.0294	0.2581
$\widehat{U}_i^{(3)}(t)$	0.0327	0.2605
$\widehat{Y}_i^{(1)}(t)$	0.0292	0.2232
$\widehat{Y}_i^{(2)}(t)$	0.0247	0.1923
$\widehat{Y}_i^{(3)}(t)$	0.0349	0.2146
	MSE	
$\widehat{\alpha}_1$	0.1061	0.1080
$\widehat{\alpha}_2$	0.0741	0.0811
$\widehat{\alpha}_3$	0.0060	0.0801
$\widehat{\nu}$	0.0054	0.0054
$\widehat{\sigma}^{2(1)}$	0.0021	0.0024
$\widehat{\sigma}^{2(2)}$	0.0023	0.0025
$\widehat{\sigma}^{2(3)}$	0.0028	0.0028

Table B.3: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) and length of the 95% credible bands for simulation cases with aligned and misaligned time points and varying measurement error variance from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on two-dimensional multivariate response with $n = 49$ regions and 200 Monte Carlo runs.

Time points	**Aligned**						**Misaligned**					
Noise level	0.2			2			0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
						MSDE						
$\widehat{\beta}_0^{(1)}(t)$	0.6173	0.3292	0.0033	0.6258	0.3584	0.0178	0.6159	0.3759	0.0130	0.6226	0.4843	0.0590
$\widehat{\beta}_0^{(2)}(t)$	0.6466	0.3260	0.0031	0.6502	0.3487	0.0236	0.6770	0.3700	0.0107	0.6772	0.4912	0.0551
$\widehat{\beta}_1^{(1)}(t)$	0.0074	0.0086	0.0017	0.0139	0.0184	0.0096	0.0211	0.0052	0.0038	0.0149	0.0150	0.0095
$\widehat{\beta}_1^{(2)}(t)$	0.0027	0.0079	0.0013	0.0109	0.0199	0.0093	0.0184	0.0050	0.0036	0.0131	0.0145	0.0096
$\widehat{\beta}_2^{(1)}(t)$	0.0024	0.0040	0.0011	0.0181	0.0397	0.0126	0.0050	0.0072	0.0038	0.0156	0.0172	0.0089
$\widehat{\beta}_2^{(2)}(t)$	0.0020	0.0035	0.0009	0.0150	0.0367	0.0093	0.0041	0.0063	0.0034	0.0143	0.0150	0.0079
$\widehat{U}_i^{(1)}(t)$	/	0.5889	0.0465	/	0.6087	0.3447	/	0.6260	0.2234	/	0.8880	0.3773
$\widehat{U}_i^{(2)}(t)$	/	0.5906	0.0470	/	0.6088	0.3516	/	0.6242	0.2207	/	0.8996	0.3623
$\widehat{Y}_i^{(1)}(t)$	0.7536	0.0289	0.0251	0.7654	0.3537	0.3499	0.8235	0.0429	0.0398	0.8317	0.3104	0.2693
$\widehat{Y}_i^{(2)}(t)$	0.6079	0.0229	0.0205	0.6222	0.2207	0.2152	0.7064	0.0311	0.0278	0.7194	0.2389	0.2241
						CP (%)						
$\widehat{\beta}_0^{(1)}(t)$	21.54	100	100	29.79	98	98.50	20.13	100	100	31.02	97.50	98
$\widehat{\beta}_0^{(2)}(t)$	18.04	99.50	100	25.52	100	100	16.33	100	100	25.09	98.50	99.50
$\widehat{\beta}_1^{(1)}(t)$	99.96	100	100	99.92	98	99	99.96	100	100	99.92	99	99.50
$\widehat{\beta}_1^{(2)}(t)$	100	100	100	100	100	100	100	100	100	99.92	99.50	99.50
$\widehat{\beta}_2^{(1)}(t)$	100	100	100	99.94	100	100	100	100	100	99.94	100	100
$\widehat{\beta}_2^{(2)}(t)$	100	100	100	100	100	100	100	100	100	100	100	100
						Length						
$\widehat{\beta}_0^{(1)}(t)$	0.554	3.446	1.008	0.896	2.167	1.067	0.608	3.520	1.223	1.002	4.205	1.767
$\widehat{\beta}_0^{(2)}(t)$	0.579	4.177	0.900	0.945	2.946	1.335	0.613	4.354	1.744	1.023	4.613	1.967
$\widehat{\beta}_1^{(1)}(t)$	0.430	0.850	0.366	0.764	1.393	0.793	0.314	0.914	0.624	0.571	1.408	0.890
$\widehat{\beta}_1^{(2)}(t)$	0.501	0.900	0.448	0.796	1.354	0.813	0.329	0.963	0.756	0.572	1.525	0.948
$\widehat{\beta}_2^{(1)}(t)$	0.402	0.756	0.397	0.712	1.770	0.994	0.430	0.902	0.670	0.622	1.489	0.943
$\widehat{\beta}_2^{(2)}(t)$	0.405	0.678	0.386	0.714	1.704	0.876	0.346	0.844	0.633	0.560	1.448	0.877

Table B.4: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) of the 95% credible bands and the corresponding length when $n = 49$ regions from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on three-dimensional multivariate response with aligned time points and 200 Monte Carlo runs.

Number of regions, n	**49 regions**					
Noise level, σ^2	0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
			MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.4664	0.0155	0.0047	0.6806	0.1079	0.0576
$\widehat{\beta}_0^{(2)}(t)$	0.5008	0.0304	0.0059	0.6808	0.1423	0.0651
$\widehat{\beta}_0^{(3)}(t)$	0.4566	0.0189	0.0097	0.6676	0.1582	0.0582
$\widehat{\beta}_1^{(1)}(t)$	0.0142	0.0021	0.0012	0.0175	0.0125	0.0097
$\widehat{\beta}_1^{(2)}(t)$	0.0200	0.0021	0.0011	0.0202	0.0114	0.0090
$\widehat{\beta}_1^{(3)}(t)$	0.0139	0.0020	0.0011	0.0392	0.0187	0.0114
$\widehat{\beta}_2^{(1)}(t)$	0.0048	0.0014	0.0009	0.0142	0.0128	0.0092
$\widehat{\beta}_2^{(2)}(t)$	0.0089	0.0015	0.0008	0.0157	0.0144	0.0093
$\widehat{\beta}_2^{(3)}(t)$	0.0119	0.0017	0.0010	0.0145	0.0131	0.0098
$\widehat{U}_i^{(1)}(t)$	/	0.0661	0.0309	/	0.3901	0.2417
$\widehat{U}_i^{(2)}(t)$	/	0.0651	0.0294	/	0.3899	0.2581
$\widehat{U}_i^{(3)}(t)$	/	0.0711	0.0327	/	0.4155	0.2605
$\widehat{Y}_i^{(1)}(t)$	0.2691	0.0386	0.0292	0.3365	0.2854	0.2232
$\widehat{Y}_i^{(2)}(t)$	0.2788	0.0256	0.0247	0.3101	0.2328	0.1923
$\widehat{Y}_i^{(3)}(t)$	0.2792	0.0456	0.0349	0.3431	0.2471	0.2146
			CP (%)			
$\widehat{\beta}_0^{(1)}(t)$	16.73	99.50	99.50	33.83	100	100
$\widehat{\beta}_0^{(2)}(t)$	11.98	99.50	99.50	26.10	100	100
$\widehat{\beta}_0^{(3)}(t)$	27.00	99.50	100	33.42	100	100
$\widehat{\beta}_1^{(1)}(t)$	98.31	100	100	99.94	100	100
$\widehat{\beta}_1^{(2)}(t)$	96.50	100	100	99.98	100	100
$\widehat{\beta}_1^{(3)}(t)$	93.48	100	100	99.46	100	100
$\widehat{\beta}_2^{(1)}(t)$	99.83	100	100	99.98	100	100
$\widehat{\beta}_2^{(2)}(t)$	99.83	100	100	99.98	100	100
$\widehat{\beta}_2^{(3)}(t)$	98.48	100	100	99.98	100	100
			Length			
$\widehat{\beta}_0^{(1)}(t)$	0.526	1.062	0.668	1.028	2.974	1.971
$\widehat{\beta}_0^{(2)}(t)$	0.482	1.094	0.636	0.845	2.913	1.996
$\widehat{\beta}_0^{(3)}(t)$	0.509	0.976	0.739	0.884	2.459	1.893
$\widehat{\beta}_1^{(1)}(t)$	0.416	0.492	0.320	0.890	1.205	0.894
$\widehat{\beta}_1^{(2)}(t)$	0.490	0.499	0.315	0.934	1.118	0.892
$\widehat{\beta}_1^{(3)}(t)$	0.409	0.438	0.318	0.920	1.310	0.890
$\widehat{\beta}_2^{(1)}(t)$	0.342	0.382	0.282	0.839	1.188	0.908
$\widehat{\beta}_2^{(2)}(t)$	0.449	0.399	0.297	0.888	1.220	0.915
$\widehat{\beta}_2^{(3)}(t)$	0.341	0.420	0.301	0.848	1.173	0.919

Appendix C

Supporting Information for Chapter 3

C.1 Prior Distributions and MCMC Implementation

For both data analysis and simulation studies, normal priors with mean zero and variance $\sigma^2_\star$ were utilized in model fits for the fixed-effects parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_\ell^\mathsf{T}, \boldsymbol{\beta}_r^\mathsf{T}, \boldsymbol{\beta}_t^{(1)\mathsf{T}}, \boldsymbol{\beta}_t^{(2)\mathsf{T}})$, $\boldsymbol{\phi} = (\boldsymbol{\phi}_\ell^\mathsf{T}, \boldsymbol{\phi}_r^\mathsf{T}, \boldsymbol{\phi}_t^{(1)\mathsf{T}}, \boldsymbol{\phi}_t^{(2)\mathsf{T}})$, γ, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, and the association parameters $\boldsymbol{\eta} = (\eta_{r0}, \eta_{r1}, \eta_t^{(1)}, \eta_t^{(2)})$ and $\boldsymbol{\zeta} = (\zeta^{(1)}, \zeta^{(2)})$, where the hyperparameters were drawn from $\sigma^2_\star \sim \text{IG}(1, 0.005)$. For the variance terms, $\boldsymbol{\sigma}^2 = (\sigma^2_{b_0}, \sigma^2_{b_1}, \sigma^2_\varepsilon, \sigma^2_\nu)$, IG priors with parameters $(a_{1*} = 0.001, a_{2*} = 0.001)$ were utilized ($*$ denoting b_0, b_1, ε or ν). We conducted sensitivity analyses with respect to hyperparameters and variance components in the current and a previous similar studies (Kürüm et al., 2024) which demonstrated insensitivity to changes in the hyperparameter settings.

The baseline hazard functions $h_{r0}(g_{ij})$, $h_{t0}^{(1)}(t)$ and $h_{t0}^{(2)}(t)$ were estimated via Bayesian P-splines with 10 equally-spaced knots and a second-order penalty. IG priors with parameters $(a_{1\kappa} = 1, a_{2\kappa} = 0.005)$ were utilized in estimation of the variance parameters κ_{h_0} associated with the baseline hazard functions. With respect to the number of knots used for the baseline hazard, we followed the results and recommendations provided by Eilers and Marx (1996) and Lang and Brezger (2004) and chose 10 equally-spaced knots while estimating the baseline hazard functions using P-splines. These authors demonstrated via extensive simulation studies that equally-spaced knots with a choice of a moderately large number of knots should be enough to ensure adequate flexibility for various different functional forms. In addition, under the P-splines, Ruppert (2002) found that the number of knots is not a crucial parameter since the smoothing is controlled by the penalty.

All computations were performed in R (version 4.0.2). Since closed-form solutions for the posterior distributions are unavailable, we fitted the proposed BM-JM and comparative models using the Bayesian software JAGS (version 4.3.0) via the `rjags` package. JAGS employs a combination of Metropolis sampling, Gibbs sampling, and other MCMC algorithms for model fitting (Plummer, 2017). For the CRIC study data analysis, three parallel chains were used, each consisting of 40,000 iterations, with an initial burn-in of 20,000 iterations. Thinning was applied to retain 2,000 posterior samples per chain, resulting in a cumulative total of 6,000 samples for subsequent estimation and inference. This process took approximately two weeks on a computing server equipped with a 64-core processor. For the simulation study, each dataset was analyzed using three parallel chains of 15,000 iterations, with an initial burn-in of 5,000 iterations. Thinning retained 2,000 posterior samples per chain, yielding a total of 6,000 samples for estimation and inference. The computation time for each dataset was approximately 20 hours. To assess the convergence of the MCMC samples for parameters associated with the longitudinal, recurrent, and competing-risk terminal event submodels in the BM-JM, trace plots are provided in Figures C.1 – C.8. Additionally, we monitored the scale reduction factor, R, as recommended by Gelman and Rubin (1992), where $R \sim 1$ indicates convergence. These diagnostics confirmed the reliability of the results for both the CRIC study data and simulation analyses.

C.2 Details on Monte Carlo Estimation of $\pi_p^{(w)}(t, s)$

To recall, the first part of the integrand of $\pi_p^{(w)}(t, s)$ is targeted via (3.6) of Section 3.3.1, while the second part is targeted using the posterior distribution of the modeling parameters based on the training dataset $\mathcal{D}_n$. The Monte Carlo simulation scheme for estimation of $\pi_p^{(w)}(t, s)$ is outlined in Table C.1, where details of each step are provided below.

Step 1: Following the Bayesian estimation proposed for BM-JM, parameters estimates are drawn from the posterior samples: $\boldsymbol{\theta}^{(h)} \sim \{\boldsymbol{\theta} \mid \mathcal{D}_n\}$ in the hth iteration, for $h = 1, \ldots, H$ (with H denoting the total number of Monte Carlo iterations utilized for dynamic predictions).

Step 2: The posterior distribution of the random effects, given the observed data, $p(\mathbf{u}_p \mid \cup_{w=1}^{2} (T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta})$ in (3.6) (see Section 3.3.1) is of nonstandard form, and thus a more sophisticated approach is required to sample from it. Note that

$$\begin{aligned} & p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}\right) \\ & \quad \propto \; P\left(\cup_{w=1}^{2}(T_p^{*(w)} > t) \mid \mathbf{u}_p; \boldsymbol{\theta}\right) \; p(\mathbf{Y}_p \mid \mathbf{b}_p; \boldsymbol{\theta}) \; p(\mathbf{G}_p, \boldsymbol{\lambda}_p \mid \mathbf{u}_p; \boldsymbol{\theta}) \; p(\mathbf{u}_p; \boldsymbol{\theta}), \quad \text{(C.1)} \end{aligned}$$

so we can use the likelihood of our data along with a Metropolis-Hastings algorithm to generate draws from $\mathbf{u}_p$ (Rizopoulos, 2011). The Metropolis-Hastings algorithm uses independent proposals $\widetilde{\mathbf{u}}_p$ from a multivariate t distribution with four degrees of freedom (MVT_4) centered at the empirical Bayesian estimates $\widehat{\mathbf{u}}_p = \text{argmax}_{\mathbf{u}} \log p(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \widehat{\boldsymbol{\theta}})$, with scale matrix

$$\widehat{\text{Var}}(\widehat{\mathbf{u}}_p) = \mathcal{I}^{-1}(\widehat{\mathbf{u}}_p) = \left\{ -\partial^2 \log p(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \widehat{\boldsymbol{\theta}}) / \; \partial \mathbf{u}_p^{\mathsf{T}} \partial \mathbf{u}_p \mid_{\mathbf{u}_p = \widehat{\mathbf{u}}_p} \right\}^{-1},$$

where $\mathcal{I}$ denotes Fisher information matrix.

The rationale for employing a MVT_4 proposal is two-fold. First, as more longitudinal measurements and recurrent event information is incorporated, the survival and random effect components in (C.1) remain relatively constant. Typically, due to the abundance of longitudinal measurements compared to limited recurrent events information, the dominant influence on the log posterior distribution of random effects is from the logarithm of the density of the linear mixed model, which tends to exhibit a roughly normal distribution. Secondly, for subjects with a limited number of longitudinal measurements and recurrent events, the heavier tails of the t distribution ensure robust coverage (Rizopoulos, 2011). Once $\widehat{\mathbf{u}}_p$ and $\widetilde{\mathbf{u}}_p$ are obtained, we proceed to calculate the Metropolis-Hastings acceptance ratio

$$\Lambda = \min \left\{ \frac{p\left(\widetilde{\mathbf{u}}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}^{(h)}\right)}{p\left(\widehat{\mathbf{u}}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}^{(h)}\right)}, 1 \right\}.$$

If a randomly generated $\widetilde{\Lambda} \sim \text{Unif}(0, 1)$ is less than Λ, i.e., $\widetilde{\Lambda} < \Lambda$, then proposal $\widetilde{\mathbf{u}}_p$ is

accepted $\mathbf{u}_p^{(h)} \equiv \widetilde{\mathbf{u}}_p$, otherwise, $\mathbf{u}_p^{(h)} \equiv \widehat{\mathbf{u}}_p$.

Step 3: Combining (3.4) and (3.6) (see Section 3.3.1), we denote the cumulative incidence probability for the wth competing-risk terminal event conditional on $\mathbf{u}_p^{(h)}$ (from Step 2) and $\boldsymbol{\theta}^{(h)}$ (from Step 1) by $\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$, and target it via

$$\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}) = \frac{\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})}{\text{S}_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})},$$

where $\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ and $\text{S}_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ are calculated using (3.7) (see Section 3.3.1) by plugging in $\mathbf{u}_p^{(h)}$ and $\boldsymbol{\theta}^{(h)}$. Since the integrals involved in the calculation of the $\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ and $\text{S}_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ do not have closed-form solutions, they are approximated via the 15-point Gauss-Kronrod quadrature rule (Kahaner et al., 1989).

Step 4: Steps 1-3 are repeated H times. Estimates of $\pi_p^{(w)}(t, s)$ are obtained as the median of the realizations of $\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}), h = 1, \ldots, H\}$. Moreover, credible intervals can be obtained using the Monte Carlo sample percentiles.

C.3 Details on Predictive Accuracy Measures

The assessment of predictive performance in joint models typically revolves around calibration, assessing how effectively the model predicts observed data, and discrimination, measuring the model's ability to discriminate between patients who will experience the terminal event from those who will not (Zheng et al., 2012; Blanche et al., 2013; Schoop et al., 2011, 2008). We assess the prediction performance of the proposed BM-JM from both perspectives using two well-established accuracy measures: the area under the receiver operating characteristic curve (AUC) and the expected Brier score (BS). While AUC assesses the model's overall discrimination ability, BS quantifies the bias between predicted and observed event risk. Adapted definitions of the two measures are considered for the proposed BM-JM, as outlined below, to account for the dynamic nature of the predictions in a competing risks setting.

Let $T_p^* = \min(T_p^{*(1)}, T_p^{*(2)})$ denote the true event time for the new patient p, with $T_p =$

$\min(T_p^*, C_p)$, C_p and δ_p denoting the terminal event time, censoring time and event indicator, respectively. We assume an independent and identically distributed (i.i.d.) testing sample of n_p subjects $\{(T_p, \delta_p, \pi_p^{(w)}(t,s)), p = 1, \ldots, n_p\}$, where $\pi_p^{(w)}(t,s)$ denotes a subject-p-specific prediction processes computed for varying prediction times t and prediction time windows Δt, such that $s = t + \Delta t$. Without loss of generality, we set $\pi_p^{(w)}(t,s) = 0$ for all subjects no longer at risk at prediction time t.

Based on the idea of discrimination, that a prediction tool should assign higher predicted risks of the event to subjects more likely to experience it, compared to those less likely, Blanche et al. (2015) extended AUC for dynamic prediction at time t for the wth competing-risk event

$$\text{AUC}^{(w)}(t,s) = P\left(\pi_p^{(w)}(t,s) > \pi_{p'}^{(w)}(t,s) \mid D_p^{(w)}(t,s) = 1, D_{p'}^{(w)}(t,s) = 0, T_p^* > t, T_{p'}^* > t\right),$$

where subject p and p' represent two randomly selected subjects from the testing dataset ($p, p' = 1, \ldots, n_p$ and $p \neq p'$). The indicator $D_p^{(w)}(t,s) = I_{(t<T_p^* \leq s, \delta_p = w)}$ specifies whether subject p experiences the w-th competing-risk terminal event within the time interval $(t,s]$. Therefore, for any subject p at risk at time t, $D_p^{(w)}(t,s) = 1$ if subject p experiences the w-th competing-risk terminal event within $(t,s]$, and $D_p^{(w)}(t,s) = 0$ if subject p either experiences a competing terminal event other than w during $(t,s]$ or remains event-free at time s. The dynamic AUC can be interpreted as the conditional probability that, for a randomly chosen pair of subjects where one experienced the event and the other did not, the predictive procedure derived from the joint model correctly ranks them.

To assess the accuracy and calibration of the prediction tool, Blanche et al. (2015) also formulated the definition of the expected dynamic Brier score (BS) for the wth competing risk $\text{BS}^{(w)}(t,s) = \mathbb{E}\left[\left\{D^{(w)}(t,s) - \pi^{(w)}(t,s)\right\}^2 \mid T^* > t\right]$, which is essentially a mean squared error measuring the average discrepancies between the true disease status and the predictive values derived from the joint model. Larger differences between probabilistic predictions and true disease status indicate more error in predictions, leading to a higher BS. Since all squared errors lie between 0 and 1, dynamic BS values are bound between 0 and 1, where a score of

0 represents perfect accuracy and a score of 1 represents perfect inaccuracy. While dynamic AUC emphasizes discrimination, focusing on the rank of predictions, dynamic BS is more oriented towards calibration, assessing the accuracy of probabilistic predictions. Therefore, the two measures complement each other in evaluating the effectiveness of the proposed dynamic prediction procedure.

For right-censored data, such as in the CRIC study, where the indicator $D_p^{(w)}(t,s)$ cannot be computed for subjects censored within $(t,s]$, Blanche et al. (2013, 2015) proposed using the Inverse Probability of Censoring Weighting (IPCW) to estimate the dynamic AUC and BS,

$$\widehat{\mathrm{AUC}}^{(w)}(t,s) = \frac{\sum_{p=1}^{n_p}\sum_{p'=1}^{n_p} I_{\{\pi_p^{(w)}(t,s)>\pi_{p'}^{(w)}(t,s)\}}\widetilde{D}_p^{(w)}(t,s)\left\{1-\widetilde{D}_{p'}^{(w)}(t,s)\right\}\widehat{W}_p(t,s)\widehat{W}_{p'}(t,s)}{\sum_{p=1}^{n_p}\sum_{p'=1}^{n_p}\widetilde{D}_p^{(w)}(t,s)\left\{1-\widetilde{D}_{p'}^{(w)}(t,s)\right\}\widehat{W}_p(t,s)\widehat{W}_{p'}(t,s)},$$

and

$$\widehat{\mathrm{BS}}^{(w)}(t,s) = \frac{1}{n_p\widehat{S}_T(t)}\sum_{p=1}^{n_p}\widehat{W}_p(t,s)\left\{\widetilde{D}_p^{(w)}(t,s)-\widehat{\pi}_p^{(w)}(t,s)\right\}^2,$$

where $\widehat{S}_T(t) = \frac{1}{n_p}\sum_{p=1}^{n_p} I_{(T_p>t)}$ estimates the probability of observing a subject at risk at time t. Note that for right-censored data, as in our CRIC study, the indicator $D_p^{(w)}(t,s)$ cannot be observed for censored subjects. To address this, we define $\widetilde{D}_p^{(w)}(t,s)$ to account for the indicator obtained only from those subjects whose outcomes are observed and known. Specifically, $\widetilde{D}_p^{(w)}(t,s) = I_{(t<T_p\le s,\delta_p=w)}$ equals 1 when subject p is known to have experienced the w-th competing-risk terminal event within $(t,s]$, and 0 otherwise.

This substitution is facilitated by IPCW using the weights $\widehat{W}_p(t,s) = I_{(T_p>s)}/\widehat{G}(s \mid t) + \{I_{(t< T_p\le s)}I_{(T_p^*\le C_p)}\}/\widehat{G}(T_p \mid t)$, where $\widehat{G}(s \mid t)$ and $\widehat{G}(T_p \mid t)$ denote the Kaplan-Meier estimators of $P(C_p > s|C_p > t)$ and $P(C_p > T_p|C_p > t)$, respectively. These represent the conditional probabilities of not being censored at time s or T_p, given that the subject was uncensored at time t. While the first term of $\widehat{W}_p(t,s)$ captures the probability of a control, a subject who experiences the terminal event after the follow-up time horizon s, among all subjects not censored at time s, the second term captures the probability of a case, a subject who experiences the terminal event within $(t,s]$ among those whose event time is not cen-

sored. By applying the weights $\widehat{W}_p(t,s)$, IPCW focuses on subjects known to be cases and controls, filtering out right-censored subjects. Consequently, IPCW leverages the observed cases and controls while enhancing their influence by incorporating their probability of being observed. Importantly, the use of IPCW in evaluating $\widehat{\text{AUC}}^{(w)}(t,s)$ and $\widehat{\text{BS}}^{(w)}(t,s)$ does not require the censoring time to be unconditionally independent of the event times. Instead, it assumes that the censoring mechanism is conditionally independent of the event times given the covariates. This conditional independence assumption ensures that the censoring mechanism is ignorable and allows the IPCW method to appropriately account for the censored observations (Robins and Rotnitzky, 1992; Uno et al., 2007; Gerds and Schumacher, 2006).

C.4 Details on Comparative Models

As mentioned in Section 3.4.3 of the main paper, we compared the proposed BM-JM with other well-established models that overlook the joint outcomes structure. Our proposed BM-JM consists of three components: the longitudinal, recurrent and competing-risk terminal components. We decompose these components into three separate marginal models, each addressing a single outcome: the longitudinal model (L-M), the recurrent model (R-M) and the competing-risk terminal model (T-M). Additionally, we combine two of these components to form simplified joint models. Since one of our primary objectives is to make dynamic predictions for competing-risk terminal outcomes, we focus on two simplified joint models: (1) the joint model of longitudinal measurements and competing-risk terminal events (referred to as LT-JM), which excludes the recurrent component, and (2) the joint model of recurrent and competing-risk terminal events (referred to as RT-JM), which excludes the longitudinal component. The detailed formulations of these comparative models are provided below.

(1) Three marginal models for longitudinal ($Y_i(t)$), recurrent ($r_{ij}(g_{ij})$), and competing-risk terminal ($h_i^{(w)}(t)$, $w = 1, 2$) outcomes are formulated as follows:

- Longitudinal model (L-M):

$$Y_i(t) = \mathbf{X}_i^{\mathsf{T}}\boldsymbol{\beta}_\ell + \mathbf{Z}_i^{\mathsf{T}}\boldsymbol{\phi}_\ell + \gamma t + b_{i0} + b_{i1}t + \varepsilon_i(t).$$

- Recurrent model (R-M):

$$r_{ij}(g_{ij}) = h_{r0}(g_{ij}) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_r + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_{rj} + \sum_{m=0}^{j-1}\alpha_m + \nu_i\right).$$

- Competing-risk terminal model (T-M):

$$\begin{aligned} h_i^{(1)}(t) &= h_{t0}^{(1)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(1)} + \nu_i\right) \\ h_i^{(2)}(t) &= h_{t0}^{(2)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(2)} + \zeta\nu_i\right). \end{aligned}$$

Note that the L-M and T-M are parameterized using follow-up time, indexed by t, whereas the R-M uses gap time, indexed by g_{ij}, which denotes the time interval between the $(j-1)$-th and j-th recurrent events for the i-th subject.

(2) Two simplified joint models are formulated as follows:

- Joint model of longitudinal measurements and competing-risk terminal events (LT-JM):

$$\begin{aligned} Y_i(t) &= \xi_i(t) + \varepsilon_i(t) = \mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_\ell + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_\ell + \gamma t + b_{i0} + b_{i1}t + \varepsilon_i(t), \\ h_i^{(1)}(t) &= h_{t0}^{(1)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(1)} + \eta_t^{(1)}\xi_i(t)\right), \\ h_i^{(2)}(t) &= h_{t0}^{(2)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(2)} + \eta_t^{(2)}\xi_i(t)\right). \end{aligned}$$

- Joint model of recurrent and competing-risk terminal events (RT-JM):

$$\begin{aligned} r_{ij}(g_{ij}) &= h_{r0}(g_{ij}) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_r + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_{rj} + \sum_{m=0}^{j-1}\alpha_m + \nu_i\right), \\ h_i^{(1)}(t) &= h_{t0}^{(1)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^\mathsf{T}\boldsymbol{\phi}_t^{(1)} + \zeta^{(1)}\nu_i\right), \\ h_i^{(2)}(t) &= h_{t0}^{(2)}(t) \exp\left(\mathbf{X}_i^\mathsf{T}\boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^{\boldsymbol{\phi}_t^{(2)} + \zeta^{(2)}\nu_i}\right). \end{aligned}$$

Note that in the LT-JM and RT-JM, the longitudinal and competing-risk terminal sub-models are parameterized using follow-up time, indexed by t, whereas the recurrent event

submodel uses gap time, indexed by g_{ij}, which denotes the time interval between the $(j-1)$-th and j-th recurrent events for the i-th subject.

C.5 Details on Simulation Studies

Simulation studies were conducted to evaluate the performance of the proposed estimation and dynamic prediction procedures, as well as to compare the proposed BM-JM with three marginal models (i.e., L-M, R-M, and T-M) and two simplified joint models (i.e., LT-JM and RT-JM) detailed in Appendix C.4. Specifically, data were generated according to the following multivariate joint model.

- Longitudinal processes: $Y_i(t) = X_i\beta_\ell + Z_i\phi_\ell + \gamma t + b_{i0} + b_{i1}t + \varepsilon_i(t)$, where $(\beta_\ell, \phi_\ell, \gamma) = (0.60, -1.30, 4)$, random (intercept and slope) effects $\mathbf{b}_i = (b_{i0}, b_{i1})^\mathsf{T} \sim N(\mathbf{0}, \mathbf{\Sigma}_b)$ with $\mathbf{\Sigma}_b = [\sigma^2_{b0}, \sigma_{01}; \sigma_{01}, \sigma^2_{b1}]$, $\sigma_{01} = \rho_b\sigma_{b0}\sigma_{b1}$, $\rho_b = 0.50$, $\sigma^2_{b0} = 1.25$, and $\sigma^2_{b1} = 0.80$, and the random error $\varepsilon_i(t) \sim N(0, \sigma^2_\varepsilon)$ with $\sigma^2_\varepsilon = 1.32$.

- Recurrent processes: $r_{ij}(g) = h_{r0}(g)\exp(X_i\beta_r + Z_i\phi_{rj} + \sum_{m=0}^{j-1}\alpha_m + \eta_{r_1}b_{i0} + \eta_{r_2}b_{i1} + \nu_i)$, where $\beta_r = 0.80$, event-varying effect $\phi_{r1} = 1.20$, $\phi_{r2} = -0.60$, and $\phi_{r3} = -0.93$, as well as event effects $\alpha_0 = 0$, $\alpha_1 = 0.60$, and $\alpha_2 = -0.80$. The link parameters between the longitudinal and recurrent processes are set to $\eta_{r_1} = 0.80$ and $\eta_{r_2} = 0.60$, while the frailty term ν_i that links the recurrent and competing-risk terminal processes is generated according to $\nu_i \sim N(0, \sigma^2_\nu)$ with $\sigma^2_\nu = 1.44$.

- Competing-risk terminal processes: $h_i^{(w)}(t) = h_{i0}^{(w)}(t)\exp(X_i\beta_t^{(w)} + Z_i\phi_t^{(w)} + \eta_t^{(w)}\xi_i(t) + \zeta^{(w)}\nu_i)$ $(w = 1, 2)$, where $\xi_i(t) = X_i\beta_\ell + Z_i\phi_\ell + \gamma t + b_{i0} + b_{i1}t$, $(\beta_t^{(1)}, \phi_t^{(1)}) = (-2.00, -1.50)$ and $(\beta_t^{(2)}, \phi_t^{(2)}) = (-1.50, -1.80)$. The link parameters between the longitudinal and the two competing-risk terminal processes are set to $\eta_t^{(1)} = 2$ and $\eta_t^{(2)} = 1.70$, respectively. Positive associations are induced between the recurrent and the two competing-risk terminal processes, by setting $\zeta^{(1)} = 1.20$ and $\zeta^{(2)} = 1.50$, respectively.

The baseline covariate X_i is generated from a normal distribution such that $X_i \sim N(2.50, 1)$, while the event-varying covariates Z_i are generated as binary variables with a probability

of 0.60. For each subject, 20 longitudinal measurement times are randomly selected on the interval $[0, 1]$ before censoring by competing-risk terminal process. The true event times for recurrent and competing-risk terminal events are simulated using the inverse probability integral transformation with Weibull baseline hazard functions: $h_{r0}(g) = 1.05t^{0.50}$ and $h_{t0}^{(w)}(t) = 1.50t^{0.50}$ $(w = 1, 2)$, respectively. This results in a probability of having one, two or three recurrent events of approximately 0.7, where similar to CRIC data analysis, our modeling considers only the first three recurrent events. In addition, the censoring rate for the competing-risk terminal events is about 30%.

Samples of 2050 subjects were generated in each Monte Carlo iteration, where the initial subset of $n = 2000$ subjects were served as the training sample utilized for estimation and inference. The remaining $n_p = 50$ subjects were designated as an independent testing sample, set aside for dynamic predictions, where $\pi^{(w)}(t, s)$ with $s = t + \Delta t$ is calculated over various combinations of prediction time t and prediction time window Δt. Specifically, for each simulated testing dataset, we derived $\widehat{\pi}_p^{(w)}(t, s) = \text{median}\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}),\ h = 1, \ldots, H\}$ over $H = 200$ Monte Carlo samples using the proposed Monte Carlo simulation procedure detailed in Appendix C.2 and Table C.1. During this process, varying prediction times t and prediction time windows Δt were set, i.e., t was set to 0, 0.3, 0.6, while Δt took on values in increments of 0.2, i.e., $(0.2, 0.4, 0.6)$, $(0.2, 0.4, 0.6)$ and $(0.2, 0.4)$ for $t = 0, 0.3, 0.6$, respectively. The prediction times $t = 0, 0.3, 0.6$ were selected to reflect an increasing number of repeated longitudinal measurements and recurrent events per subject, which is expected to result in a decreasing degree of shrinkage in the estimation of the random effects. To assess the accuracy of the proposed dynamic prediction algorithm, for each combination of t and Δt, $\widehat{\text{AUC}}(t, s)$ and $\widehat{\text{BS}}(t, s)$ values $(s = t + \Delta t)$ were computed for each competing-risk terminal outcome, as described in Section 3.3.2 of the main paper and Appendix C.3, and the results are summarized in Table C.9.

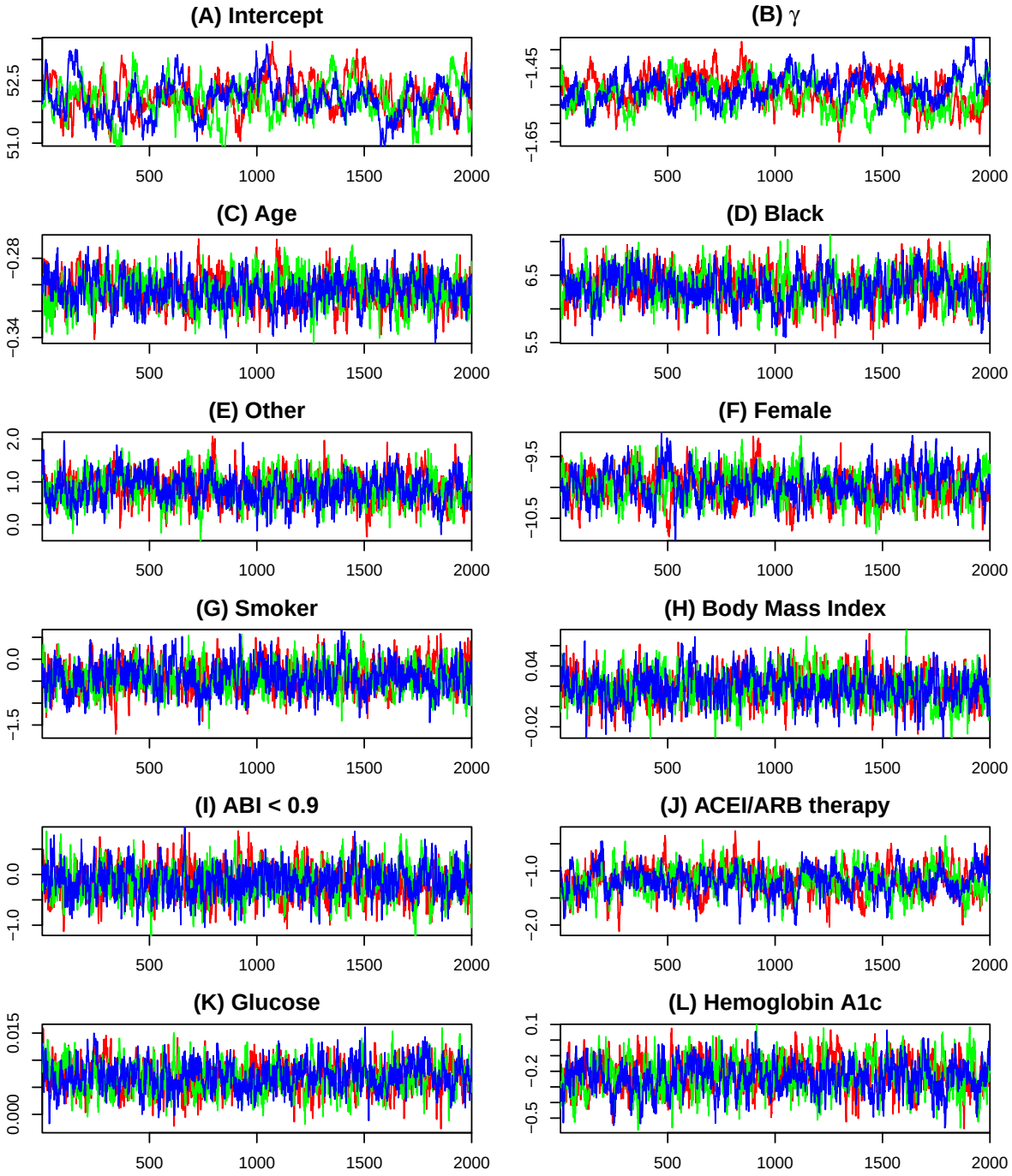

Figure C.1: Trace plots for longitudinal submodel parameters (part 1).

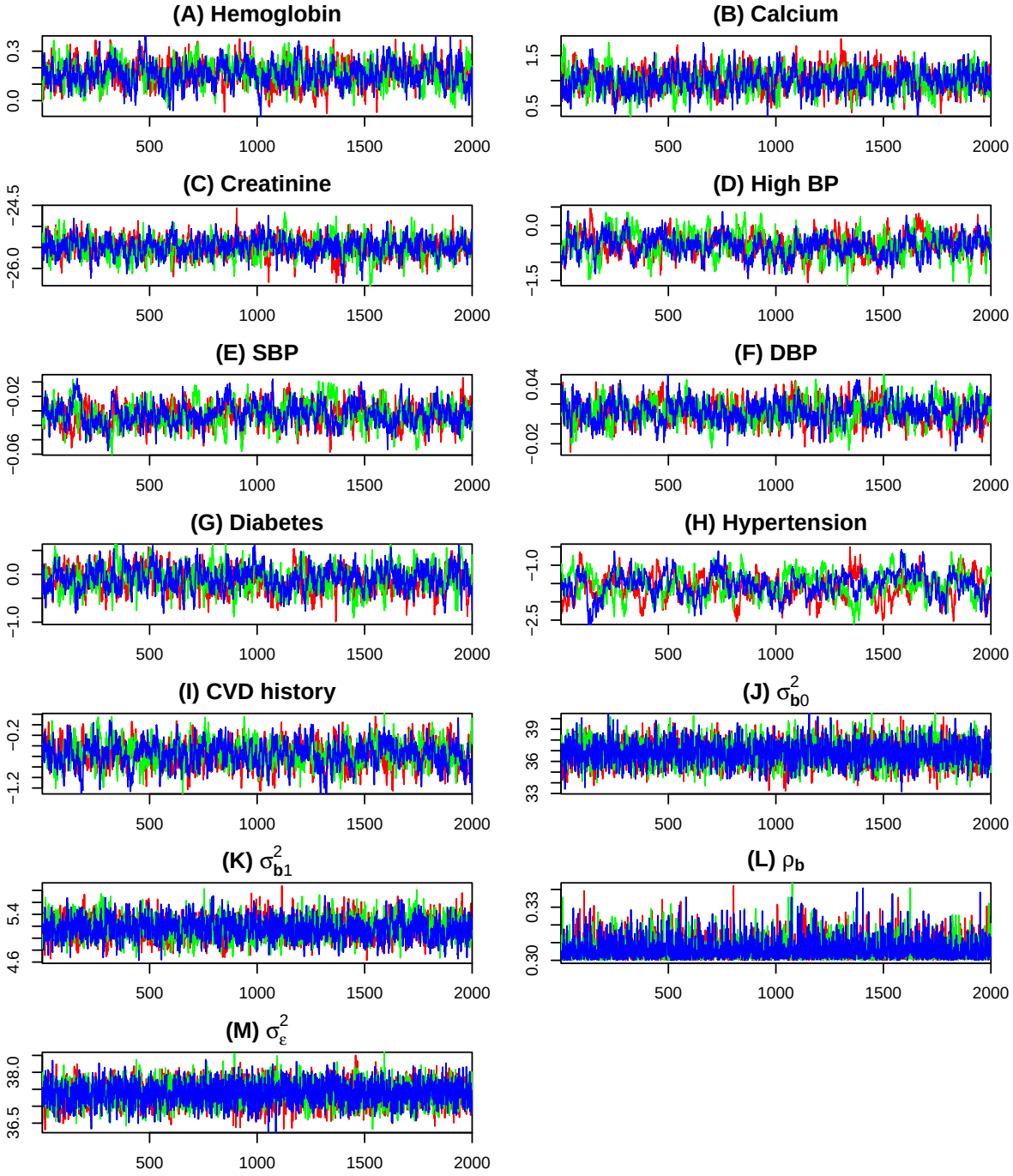

Figure C.2: Trace plots for longitudinal submodel parameters (part 2).

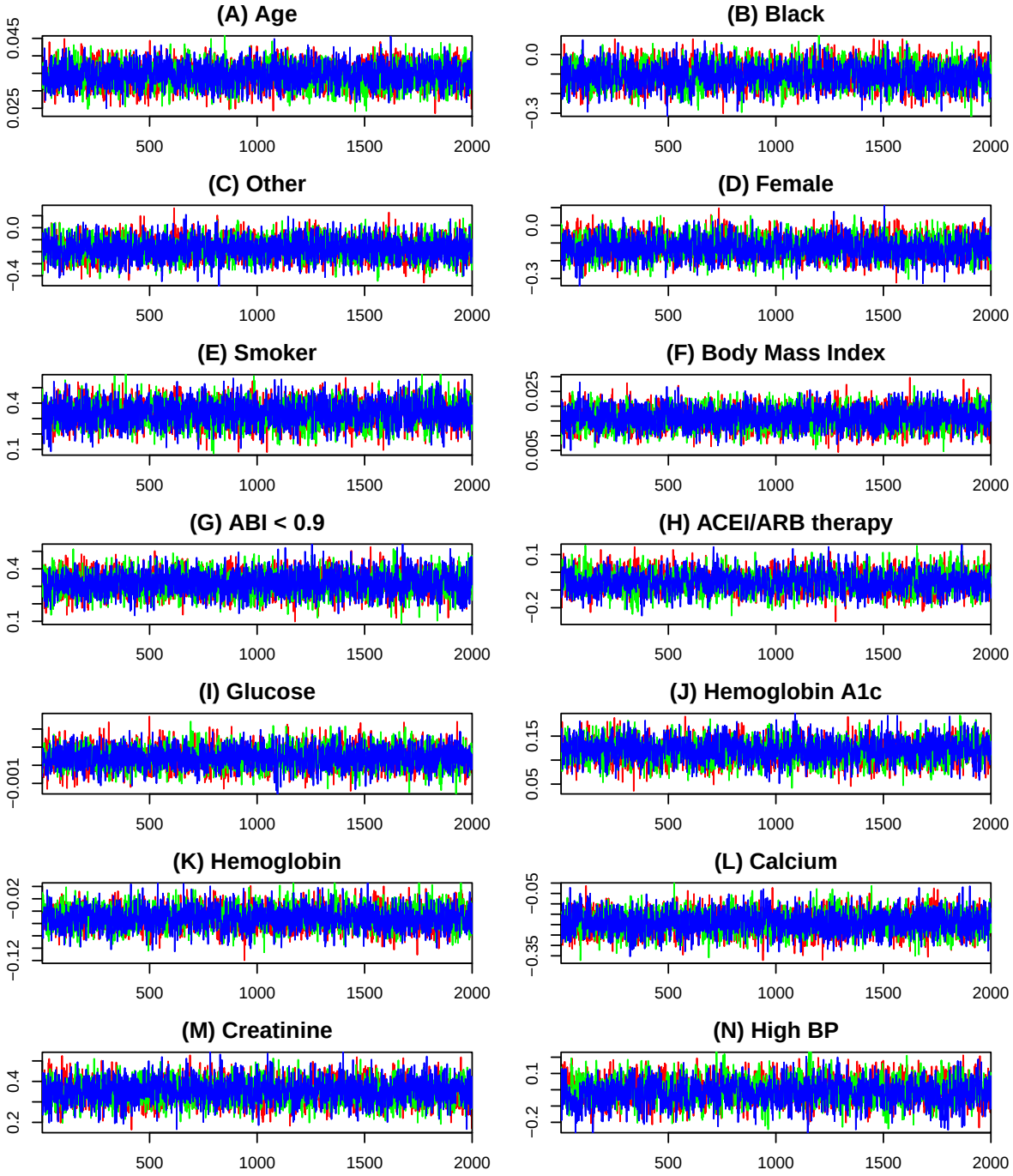

Figure C.3: Trace plots for recurrent events submodel parameters (part 1).

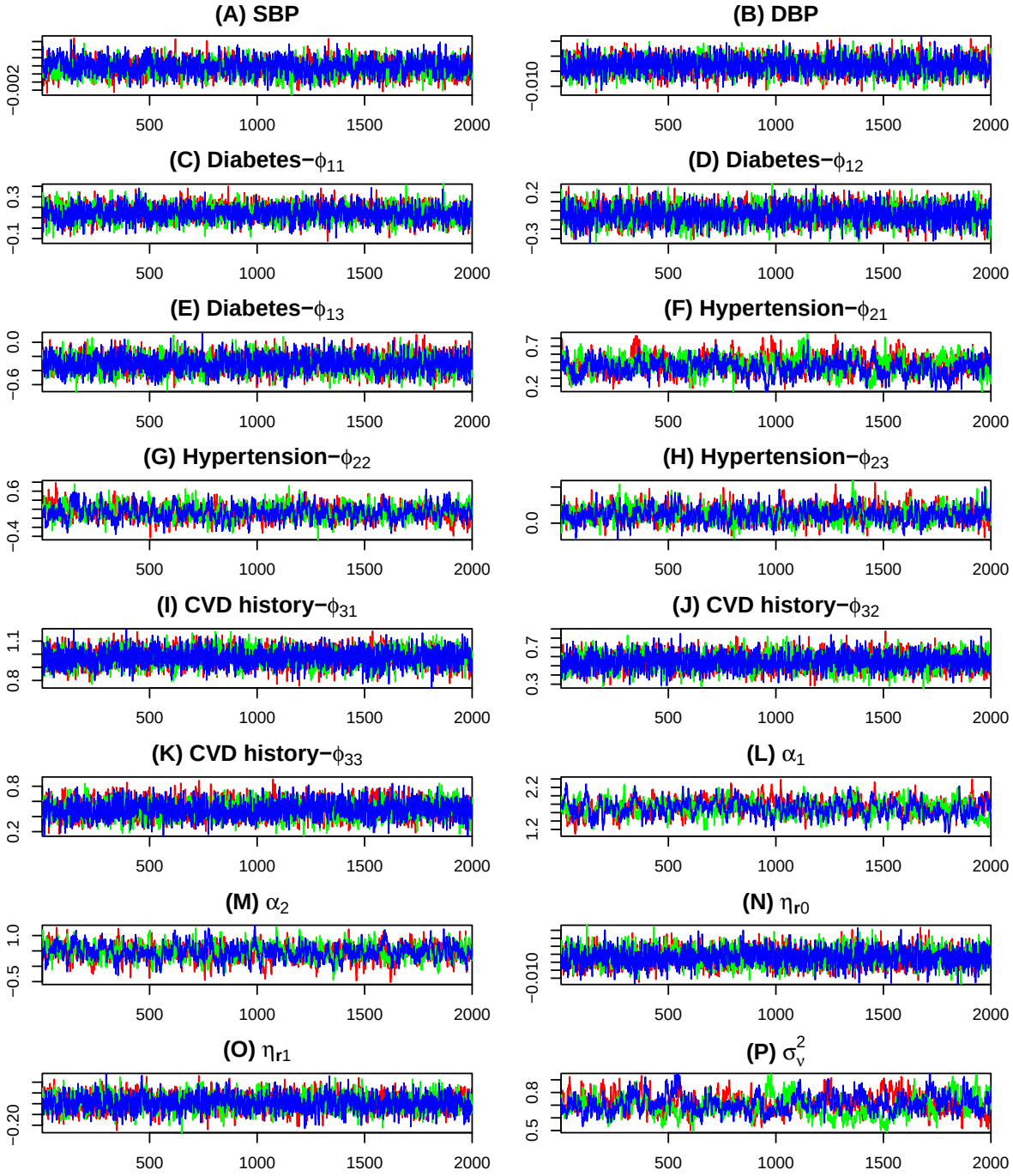

Figure C.4: Trace plots for recurrent events submodel parameters (part 2).

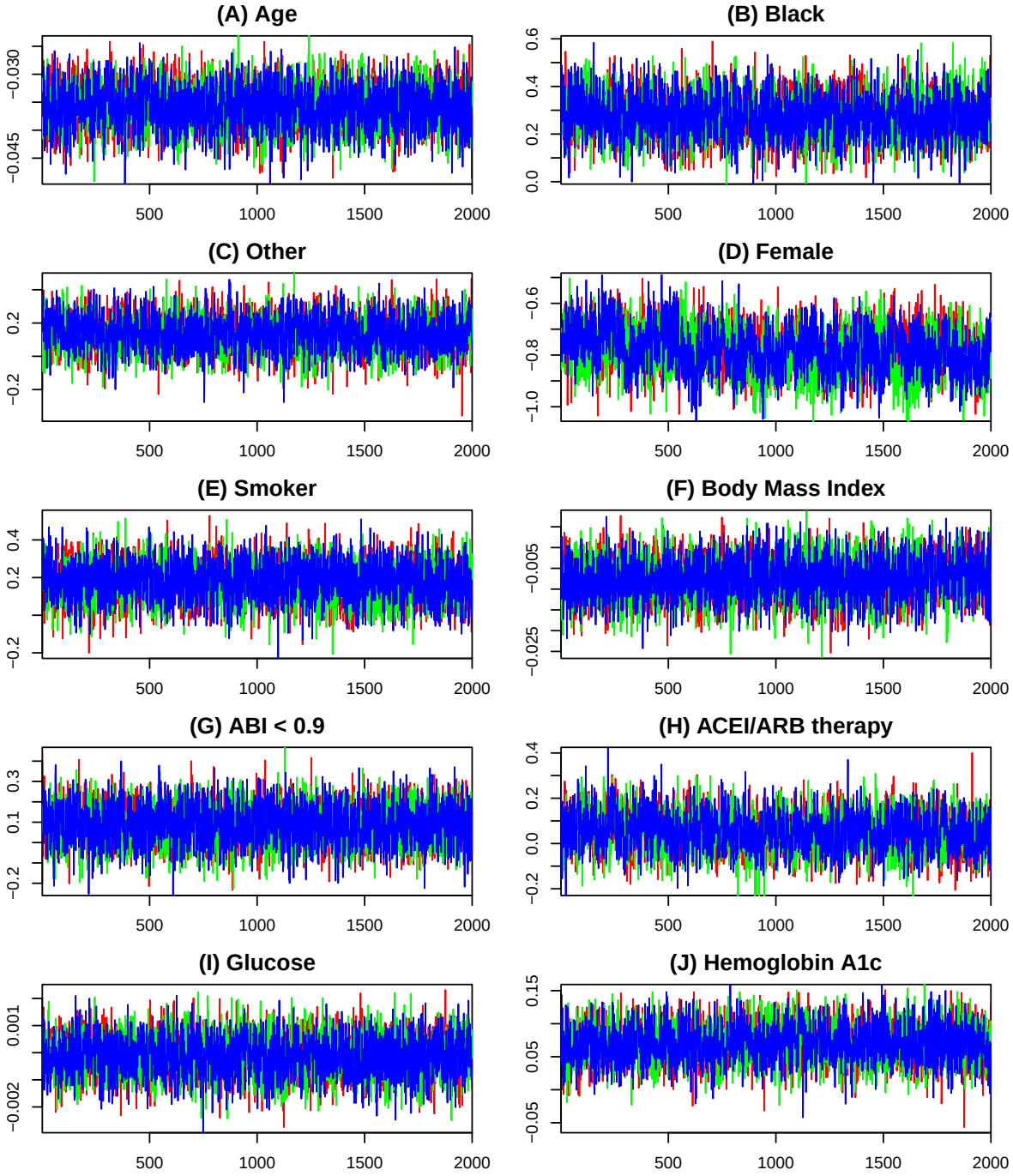

Figure C.5: Trace plots for ESKD submodel parameters (part 1).

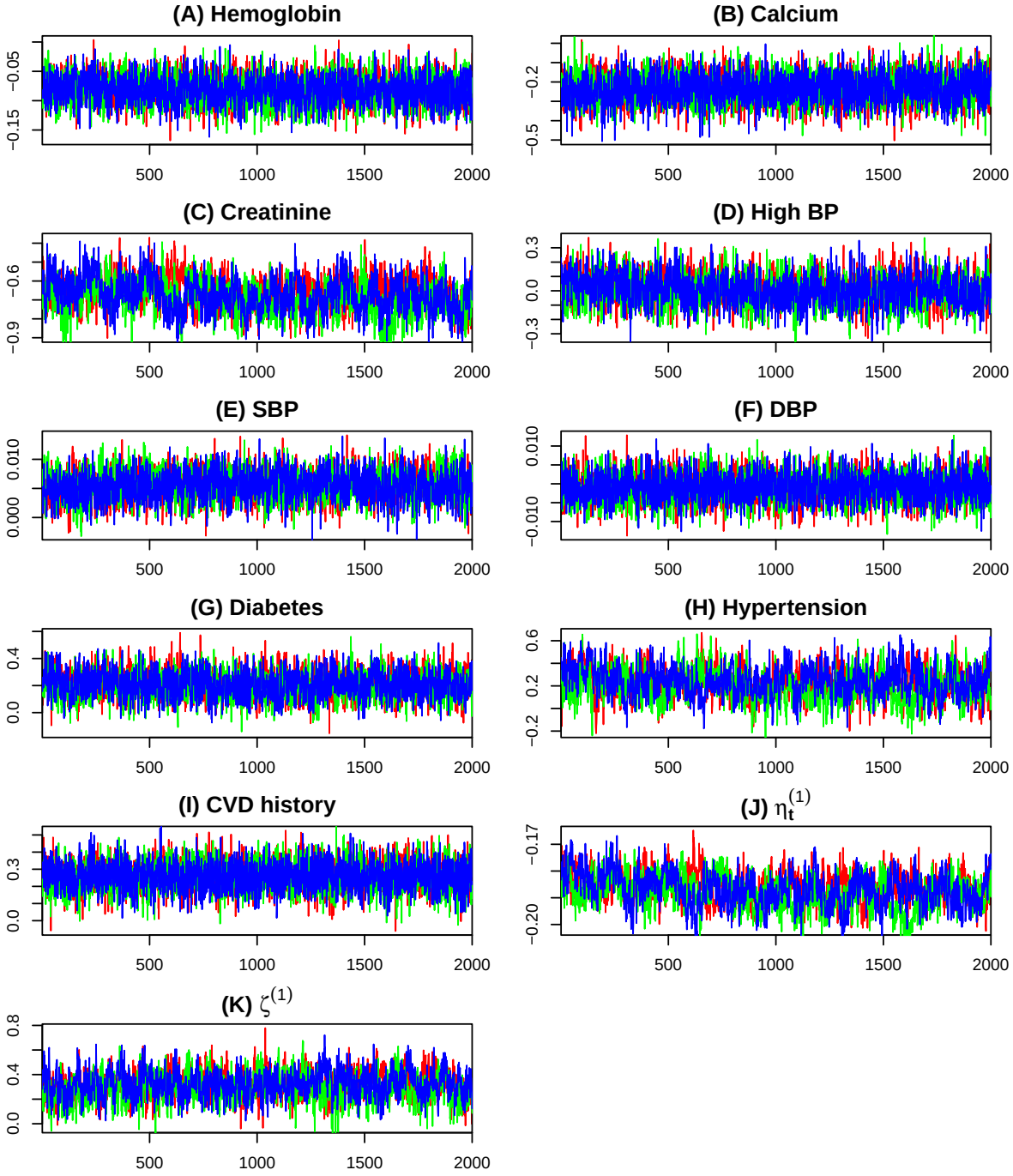

Figure C.6: Trace plots for ESKD submodel parameters (part 2).

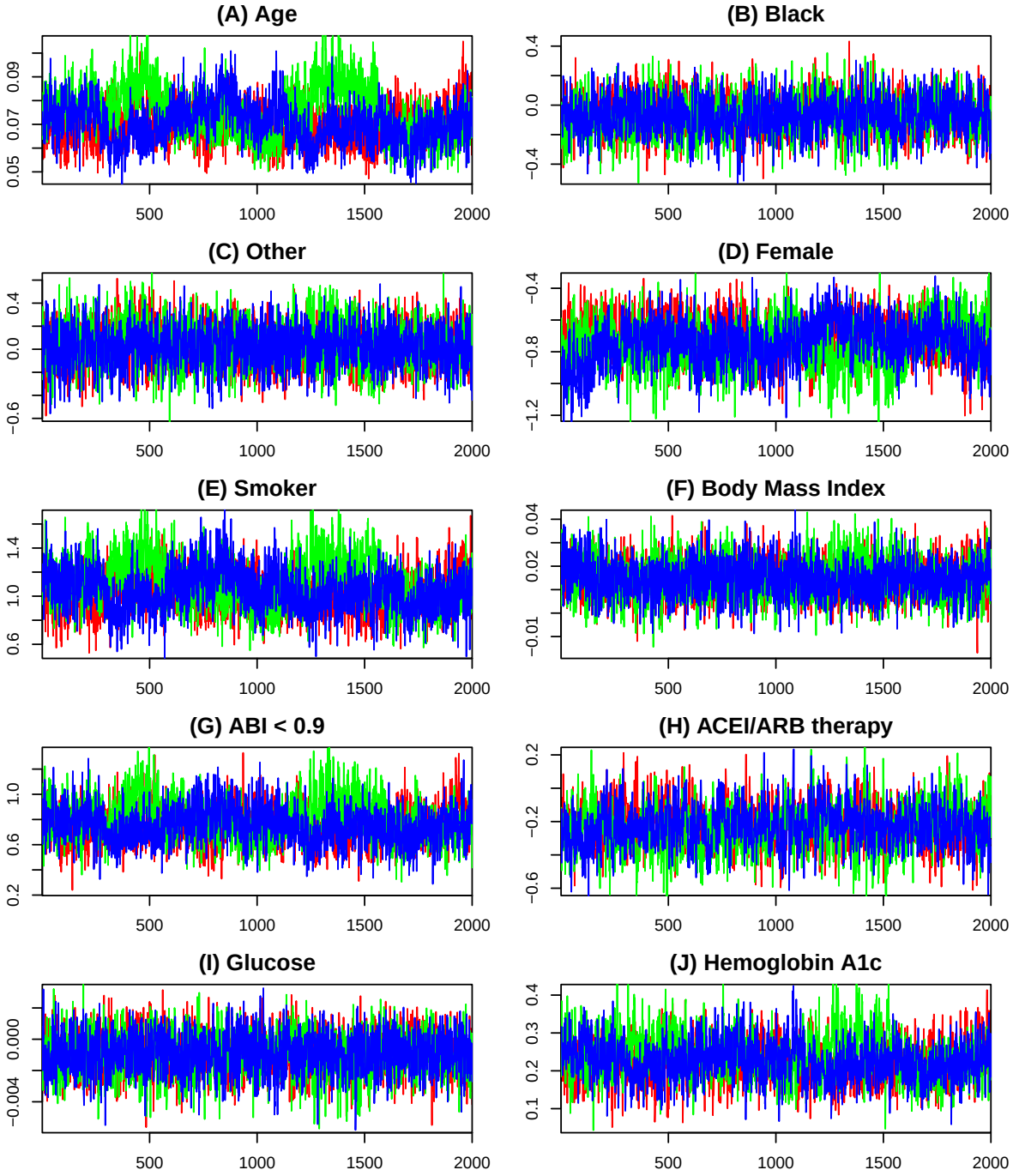

Figure C.7: Trace plots for death submodel parameters (part 1).

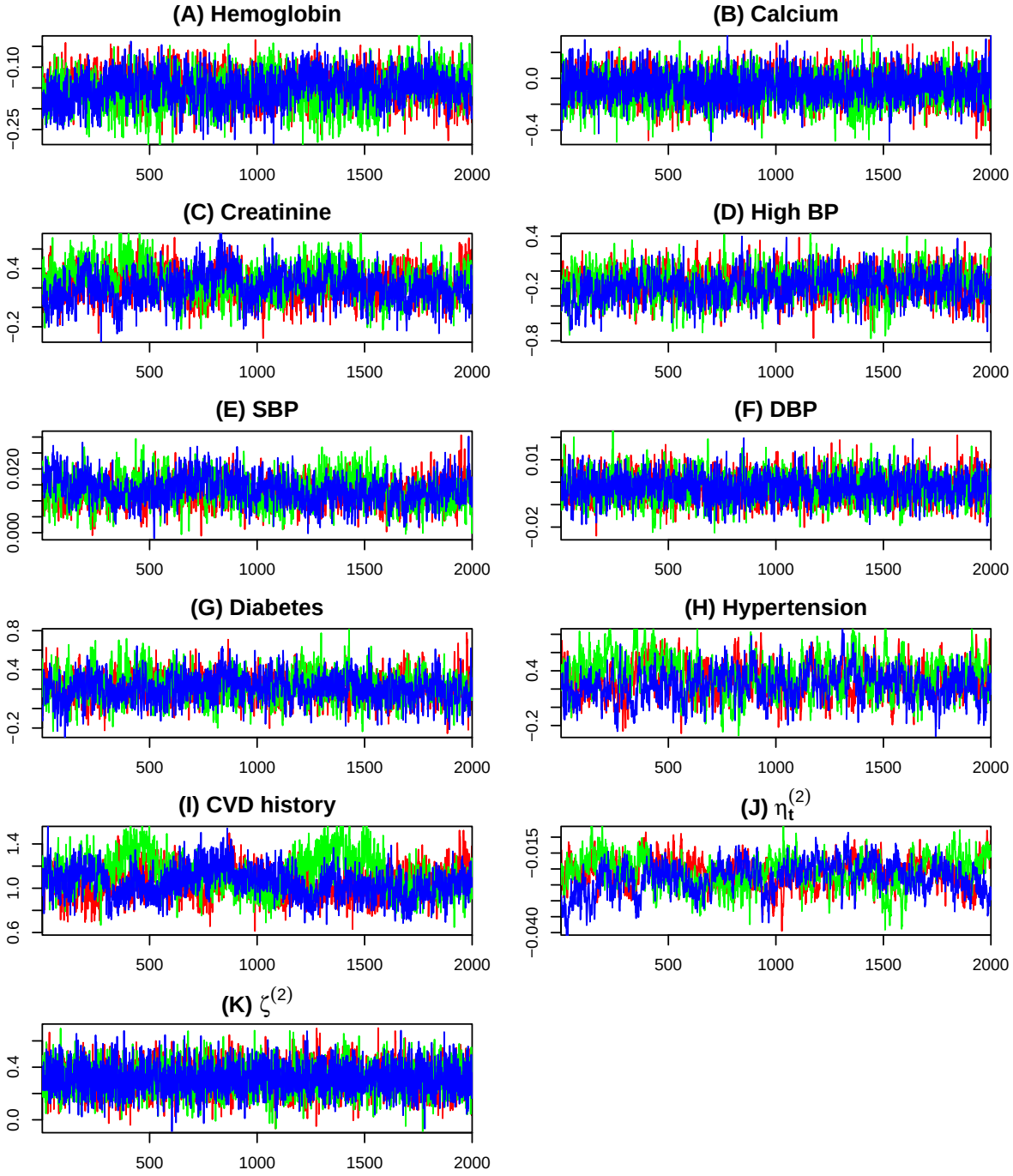

Figure C.8: Trace plots for death submodel parameters (part 2).

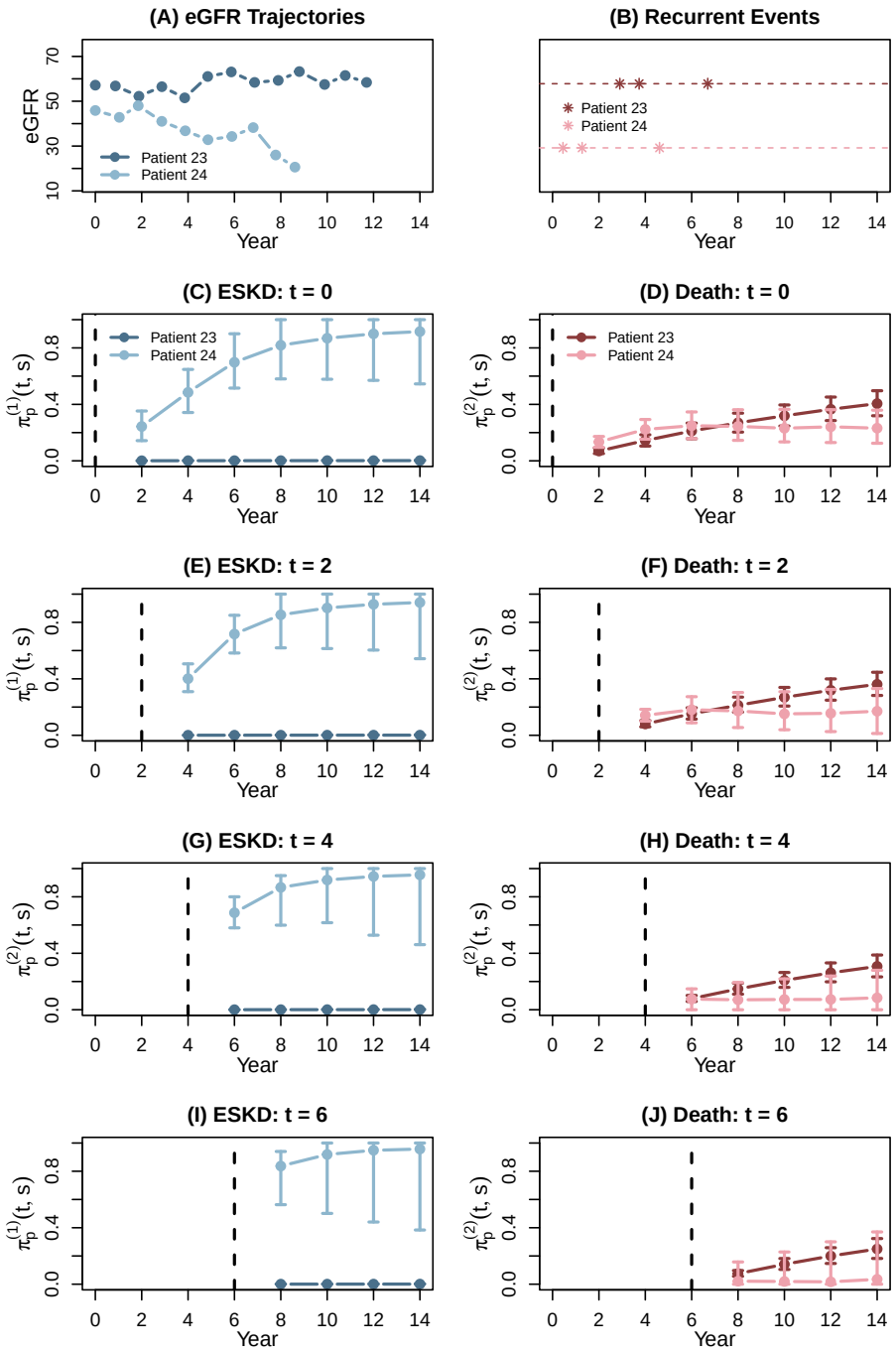

Figure C.9: The eGFR trajectories (A) and recurrent CV events profiles (B), and the estimated median $\pi_p^{(w)}(t, s)$, as well as pointwise credible intervals ((C, E, G, I) for ESKD and (D, F, H, J) for death) at different prediction times, i.e., $t = 0, 2, 4, 6$ years, for Patient 23 (dark) and Patient 24 (light). The vertical dashed line denotes the time of prediction. Note that for Patient 23, the credible intervals of ESKD prediction are too small to be visible.

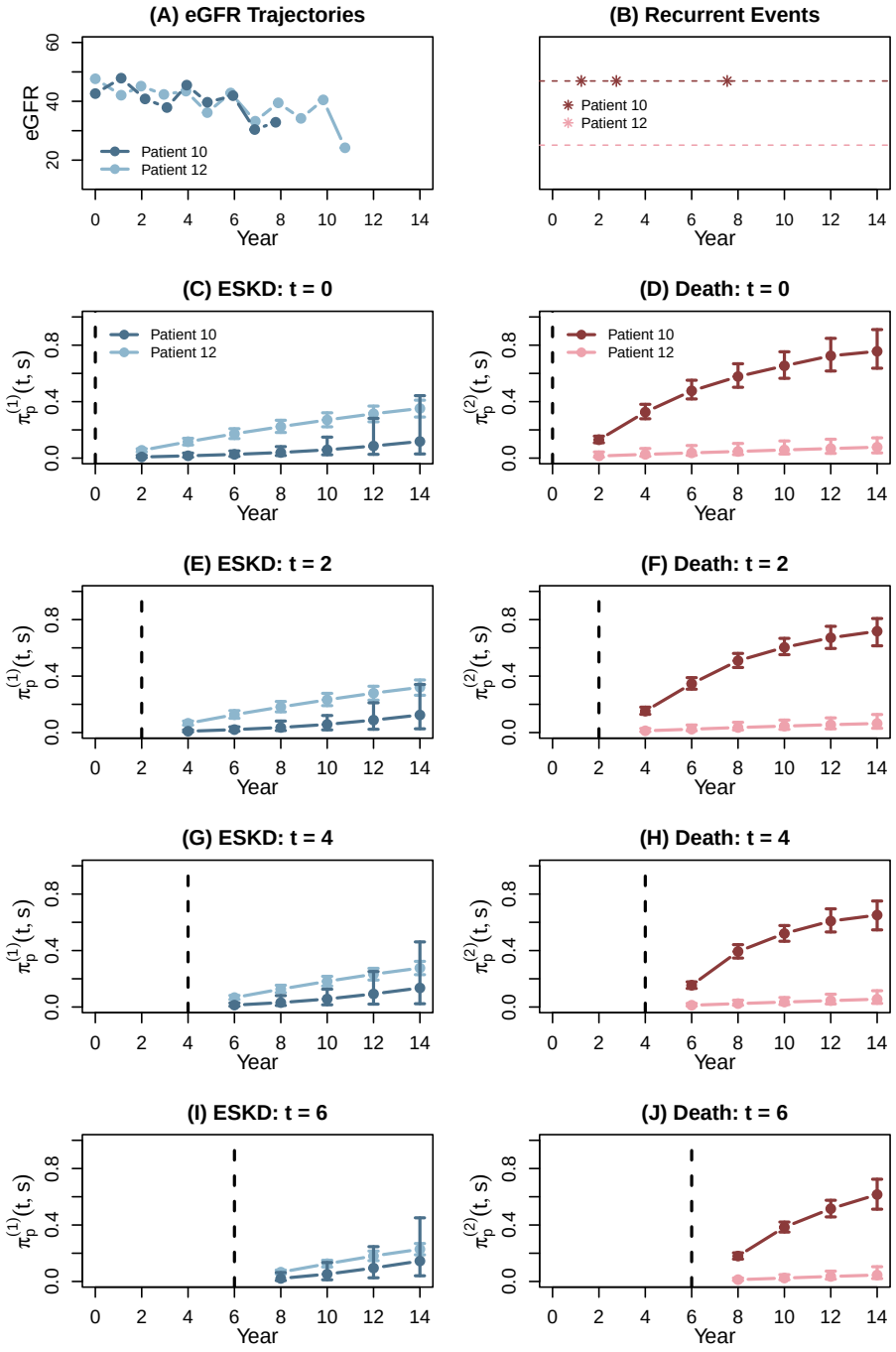

Figure C.10: The eGFR trajectories (A) and recurrent CV events profiles (B), and the estimated median $\pi_p^{(w)}(t, s)$, as well as pointwise credible intervals ((C, E, G, I) for ESKD and (D, F, H, J) for death) at different prediction times, i.e., $t = 0, 2, 4, 6$ years, for Patient 10 (dark) and Patient 12 (light). The vertical dashed line denotes the time of prediction.

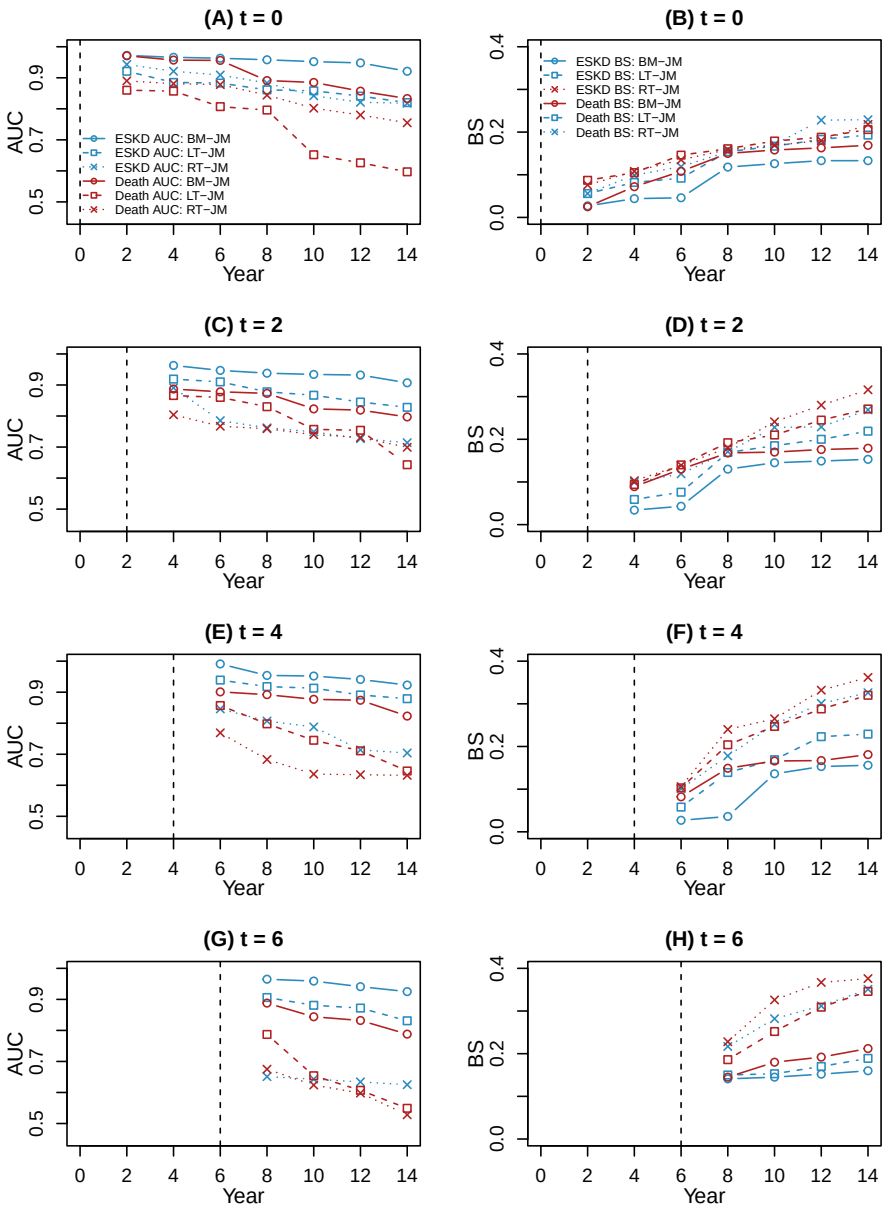

Figure C.11: Dynamic AUC (A, C, E, G) and BS (B, D, F, H) values from estimated $\pi_p^{(w)}(t, s)$ for ESKD (blue) and death (red) at different prediction times, i.e., (A, B) $t = 0$, (C, D) $t = 2$ years, (E, F) $t = 4$ years, and (G, H) $t = 6$ years, until end of follow-up. The results are based on the proposed BM-JM (circle) and simplified joint models, i.e., LT-JM (square) and RT-JM (crossing) fits on the testing set of the CRIC data. The vertical dashed line denotes the time of prediction.

Table C.1: Monte Carlo simulation procedure for drawing $\pi_p^{(w)}(t,s)$ samples under the Bayesian formulation of the proposed BM-JM.

Algorithm Dynamic Prediction Procedure

Step 1: Draw $\boldsymbol{\theta}^{(h)} \sim \{\boldsymbol{\theta} \mid \mathcal{D}_n\}$.

Step 2: Draw $\mathbf{u}_p^{(h)} \sim \left\{\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}^{(h)}\right\}$:

(a) Calculate empirical Bayes estimates $\widehat{\mathbf{u}}_p$:

$$\widehat{\mathbf{u}}_p = \arg\max_{\mathbf{u}} \; p\left(\mathbf{u}_p \mid \cup_{w=1}^{2}(T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \boldsymbol{\theta}^{(h)}\right);$$

(b) Generate proposal $\widetilde{\mathbf{u}}_p \sim \text{MVT}_4\left(\widehat{\mathbf{u}}_p, \widehat{\text{Var}}(\widehat{\mathbf{u}}_p)\right)$ where:

$$\widehat{\text{Var}}(\widehat{\mathbf{u}}_p) = \left\{ -\frac{\partial^2 \log p(\mathbf{u}_p \mid \cdots)}{\partial \mathbf{u}_p^{\mathsf{T}} \partial \mathbf{u}_p}\bigg|_{\mathbf{u}_p = \widehat{\mathbf{u}}_p} \right\}^{-1};$$

(c) Compute Metropolis-Hastings acceptance ratio:

$$\Lambda = \min\left\{\frac{p(\widetilde{\mathbf{u}}_p \mid \cdots)}{p(\widehat{\mathbf{u}}_p \mid \cdots)}, 1\right\};$$

(d) Generate $\widetilde{\Lambda} \sim \text{Unif}(0,1)$. If $\widetilde{\Lambda} < \Lambda$, accept proposal $\mathbf{u}_p^{(h)} = \widetilde{\mathbf{u}}_p$; else $\mathbf{u}_p^{(h)} = \widehat{\mathbf{u}}_p$.

Step 3: Compute

$$\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}) = \frac{\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})}{\text{S}_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})}.$$

Step 4: Repeat Steps 1-3 H times. Use the samples $\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}), h = 1, \ldots, H\}$ to estimate $\pi_p^{(w)}(t,s)$.

Table C.2: Summary of patient risk factors based on the CRIC study cohort of 5,194 individuals.

Variable	Mean (SD) or Count (Percent)*
Age (years)	59.57 (10.67)
Race/Ethnicty:	
Non-Hispanic Black	2278 (43.86)
Non-Hispanic White	2156 (41.51)
Other	760 (14.63)
Female	2252 (43.36)
Current Smoker	649 (12.50)
Body Mass Index (kg/m^2)	32.23 (7.53)
Angle-Brachial Index < 0.9	785 (15.11)
ACE Inhibitor/ARB therapy	3578 (68.89)
Estimated Glomerular Filtration Rate, baseline ($ml/min/1.73m^2$)	48.38 (15.56)
Blood glucose (mg/dL)	119.77 (54.07)
Hemoglobin A1c (%)	6.67 (1.57)
Hemoglobin (g/dL)	12.68 (1.77)
Calcium (mg/dL)	9.25 (0.51)
Creatinine (Roche adjusted, mg/dL)	1.61 (0.55)
High Blood Pressure (BP >130/80)	2456 (47.29)
Systolic Blood Pressure (mmHg)	128.19 (21.31)
Diastolic Blood Pressure (mmHg)	70.95 (12.58)
Diabetes	2656 (51.14)
Hypertension	4490 (86.45)
History of Cardiovascular Disease	1438 (27.69)

1. * For categorical variables;
2. Angiotensin-converting enzyme (ACE); angiotensin receptor blocker (ARB)

Table C.3: Results from the proposed BM-JM: (A) Estimates of associations/links among multiple outcomes: longitudinal estimated glomerular filtration rate (eGFR), recurrent cardiovascular (CV) events, competing-risk terminal events (i.e., ESKD and death), and (B) model variance components.

(A) Association/Link Parameters	Estimate	95% CI
eGFR and Recurrent CV Events Link:		
η_{r0}	0.004	(−0.006, 0.014)
η_{r1}	−0.157	(−0.187, −0.127)*
eGFR and Competing-risk Terminal Events Link:		
$\eta_t^{(1)}$	−0.188	(−0.200, −0.176)*
$\eta_t^{(2)}$	−0.021	(−0.031, −0.012)*
CV Events and Competing-risk Terminal Events Link:		
$\zeta^{(1)}$	0.285	(0.071, 0.505)*
$\zeta^{(2)}$	2.818	(2.057, 3.686)*
(B) Variance Components		
Var(eGFR Intercept), σ^2_{b0}	36.820	(34.846, 38.833)*
Var(eGFR Slope), σ^2_{b1}	5.160	(4.813, 5.527)*
Correlation(Intercept, Slope), ρ_b	0.306	(0.300, 0.321)*
Measurement Error Variance, σ^2_ϵ	37.378	(36.770, 38.013)*
σ^2_ν (Terminal & Recurrent Link Random Effect)	0.689	(0.563, 0.835)*

* 95% credible interval (CI) does not include estimate of 0

Table C.4: Results from three marginal models of (A) L-M: longitudinal estimated glomerular filtration rate (eGFR), (B) R-M: recurrent cardiovascular (CV) events, (C1 and C2) T-M: competing-risk terminal events (i.e., (C1) for ESKD and (C2) for death). Effect sizes (estimates and hazard ratios (HRs)) are given for one unit change in covariates.

	(A) Longitudinal eGFR	(B) Recurrent CV Events	(C1) Terminal ESKD Event	(C2) Terminal Death Event
Variable	Estimate (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Intercept	52.322 (51.498, 53.115)*	-	-	-
Time, γ	−1.334 (−1.408, −1.258)*	-	-	-
Age (years)	−0.305 (−0.327, −0.282)*	1.025 (1.019, 1.032)*	0.965 (0.955, 0.974)*	1.049 (1.040, 1.058)*
Non-Hispanic Black*	6.294 (5.817, 6.777)*	1.030 (0.978, 1.103)	0.868 (0.695, 1.069)	0.901 (0.779, 1.039)
Other*	0.976 (0.338, 1.627)*	0.948 (0.896, 1.013)	1.728 (1.358, 2.225)*	0.972 (0.796, 1.183)
Female	−10.004 (−10.471, −9.522)*	0.948 (0.896, 1.013)	1.306 (1.066, 1.594)*	0.761 (0.657, 0.879)*
Smoker	−0.432 (−1.064, 0.182)	1.093 (1.018, 1.195)*	1.637 (1.285, 2.087)*	1.795 (1.499, 2.121)*
Body Mass Index	0.015 (−0.014, 0.043)	1.013 (1.007, 1.020)*	0.993 (0.982, 1.004)	1.010 (1.001, 1.019)*
ABI < 0.9	−0.128 (−0.724, 0.470)	1.090 (1.014, 1.198)*	0.960 (0.762, 1.214)	1.499 (1.293, 1.731)*
ACEI/ARB use	−1.188 (−1.672, −0.724)*	0.997 (0.937, 1.063)	1.111 (0.901, 1.354)	0.904 (0.783, 1.042)
Glucose (mg/dL)	0.007 (0.002, 0.012)*	1.001 (0.999, 1.002)	1.000 (0.998, 1.002)	1.000 (0.998, 1.001)
Hemoglobin A1c (%)	−0.224 (−0.420, −0.028)*	1.110 (1.068, 1.163)*	1.184 (1.098, 1.290)*	1.136 (1.070, 1.203)*
Hemoglobin (g/dL)	0.168 (0.026, 0.310)*	0.969 (0.942, 1.015)	0.855 (0.796, 0.911)*	0.934 (0.894, 0.975)*
Calcium (mg/dL)	1.008 (0.574, 1.440)*	0.892 (0.829, 0.983)*	0.539 (0.447, 0.648)*	0.940 (0.826, 1.070)
Creatinine (mg/dL)	−25.749 (−26.174, −25.317)*	1.119 (1.040, 1.194)*	11.817 (8.544, 17.744)*	1.526 (1.317, 1.776)*
High BP	−0.616 (−1.220, −0.004)*	0.997 (0.914, 1.086)	1.334 (1.026, 1.717)*	0.937 (0.783, 1.114)
SBP (mmHg)	−0.030 (−0.046, −0.014)*	1.005 (1.000, 1.009)	1.028 (1.020, 1.037)*	1.008 (1.003, 1.013)*
DPB (mmHg)	0.015 (−0.008, 0.037)	0.995 (0.990, 1.001)	0.995 (0.986, 1.004)	0.997 (0.991, 1.003)
Diabetes	−0.098 (−0.639, 0.429)	-	1.506 (1.190, 1.915)*	1.106 (0.946, 1.298)
Diabetes - ϕ_{11}	-	1.225 (1.069, 1.414)*	-	-
Diabetes - ϕ_{12}	-	1.017 (0.843, 1.229)	-	-
Diabetes - ϕ_{13}	-	0.834 (0.660, 1.055)	-	-
Hypertension	−1.882 (−2.599, −1.162)*	-	1.268 (0.871, 1.808)	1.240 (1.002, 1.546)*
Hypertension - ϕ_{21}	-	1.677 (1.366, 2.081)*	-	-
Hypertension - ϕ_{22}	-	1.148 (0.852, 1.578)	-	-
Hypertension - ϕ_{23}	-	1.040 (0.687, 1.629)	-	-
CVD	−0.589 (−1.052, −0.124)*	-	1.389 (1.142, 1.683)*	1.886 (1.661, 2.141)*
CVD - ϕ_{31}	-	2.783 (2.440, 3.162)*	-	-
CVD - ϕ_{32}	-	1.785 (1.493, 2.137)*	-	-
CVD - ϕ_{33}	-	1.722 (1.365, 2.209)*	-	-
α_1	-	5.559 (3.813, 8.039)*	-	-
α_2	-	1.753 (1.020, 2.869)*	-	-
	Correlation/Covariance Estimate (95% CI)			
0.5pt σ^2_ϵ	37.831 (37.199, 38.494)*	-	-	-
$\sigma^2_{b_0}$	35.506 (33.532, 37.515)*	-	-	-
$\sigma^2_{b_1}$	4.437 (4.138, 4.753)*	-	-	-
ρ_b	0.311 (0.300, 0.335)*	-	-	-
σ^2_ν	-	1.786 (1.473, 2.152)*	9.467(3.879, 32.432)*	-
ζ	-	-	-	1.033(0.901, 1.215)

1. $*$ 95% credible interval (CI) does not include estimate of 0 or hazard ratio (HR) of 1;
2. $\star$ Reference group: Non-Hispanic White;
3. Angiotensin-converting enzyme inhibitor(ACEI); angiotensin receptor blocker (ARB); blood pressure (BP); cardiovascular disease (CVD); hemoglobin A1c (HbA1c); systolic and diastolic BP (SBP, DBP)

Table C.5: Results from LT-JM of (A) longitudinal estimated glomerular filtration rate (eGFR), (B1 and B2) competing-risk terminal events (i.e., (B1) for ESKD and (B2) for death). Effect sizes (estimates and hazard ratios [HRs]) are given for one unit change in covariates.

	(A) Longitudinal eGFR	(B1) Terminal ESKD Event	(B2) Terminal Death Event
Variable	Estimate (95% CI)	HR (95% CI)	HR (95% CI)
Intercept	52.018 (51.131, 52.859)*	-	-
Time, γ	−1.505 (−1.575, −1.425)*	-	-
Age (years)	−0.306 (−0.328, −0.285)*	0.964 (0.957, 0.971)*	1.046 (1.037, 1.055)*
Non-Hispanic Black*	6.317 (5.848, 6.791)*	1.325 (1.118, 1.579)*	0.931 (0.803, 1.081)
Other*	0.838 (0.201, 1.484)*	1.148 (0.936, 1.403)	0.960 (0.776, 1.179)
Female	−9.989 (−10.475, −9.532)*	0.465 (0.394, 0.546)*	0.692 (0.591, 0.810)*
Smoker	−0.402 (−1.094, 0.217)	1.186 (0.982, 1.436)	1.779 (1.493, 2.115)*
Body Mass Index	0.020 (−0.008, 0.049)	0.993 (0.983, 1.002)	1.010 (0.999, 1.019)
ABI < 0.9	−0.112 (−0.710, 0.455)	1.056 (0.877, 1.268)	1.504 (1.294, 1.740)*
ACEI/ARB use	−1.237 (−1.715, −0.748)*	1.057 (0.904, 1.235)	0.887 (0.769, 1.029)
Glucose (mg/dL)	0.007 (0.002, 0.012)*	1.000 (0.998, 1.001)	1.000 (0.998, 1.001)
Hemoglobin A1c (%)	−0.229 (−0.420, −0.033)*	1.072 (1.013, 1.132)*	1.133 (1.071, 1.199)*
Hemoglobin (g/dL)	0.170 (0.020, 0.312)*	0.926 (0.886, 0.970)*	0.937 (0.897, 0.980)*
Calcium (mg/dL)	0.992 (0.586, 1.427)*	0.799 (0.691, 0.923)*	0.969 (0.849, 1.102)
Creatinine (mg/dL)	−25.532 (−25.989, −25.088)*	0.505 (0.429, 0.589)*	1.171 (0.965, 1.417)
High BP	−0.536 (−1.192, 0.131)	1.012 (0.825, 1.239)	0.932 (0.780, 1.118)
SBP (mmHg)	−0.032 (−0.048, −0.016)*	1.005 (1.000, 1.010)	1.007 (1.002, 1.011)*
DPB (mmHg)	0.010 (−0.013, 0.034)	0.998 (0.991, 1.006)	0.997 (0.991, 1.004)
Diabetes	−0.098 (−0.568, 0.403)	1.216 (1.004, 1.466)*	1.078 (0.921, 1.265)
Hypertension	−1.453 (−2.182, −0.711)*	1.230 (0.943, 1.614)	1.217 (0.971, 1.541)
CVD	−0.534 (−1.052, −0.104)*	1.241 (1.065, 1.446)*	1.878 (1.653, 2.133)*
	Correlation/Covariance Estimate (95% CI)		
0.5pt σ^2_{b0}	36.501 (34.433, 38.691)*	-	-
σ^2_{b1}	5.188 (4.842, 5.562)*	-	-
ρ_b	0.306 (0.300, 0.322)*	-	-
σ^2_ϵ	37.406 (36.774, 38.062)*	-	-
$\eta^{(1)}_t$	-	−0.18 (−0.200, −0.176)*	-
$\eta^{(2)}_t$	-	-	−0.023 (−0.036, −0.013)*

1. * 95% credible interval (CI) does not include estimate of 0 or hazard ratio (HR) of 1;
2. * Reference group: Non-Hispanic White;

Table C.6: Results from RT-JM of (A) recurrent cardiovascular (CV) events, (B1 and B2) competing-risk terminal events (i.e., (B1) for ESKD and (B2) for death). Effect sizes (estimates and hazard ratios [HRs]) are given for one unit change in covariates.

	(A) Recurrent CV Events	(B1) Terminal ESKD Event	(B2) Terminal Death Event
Variable	HR (95% CI)	HR (95% CI)	HR (95% CI)
Age (years)	1.027 (1.021, 1.034)*	0.976 (0.968, 0.984)*	1.057 (1.045, 1.069)*
Non-Hispanic Black*	0.954 (0.854, 1.066)	0.860 (0.719, 1.031)	0.881 (0.724, 1.071)
Other*	0.914 (0.783, 1.069)*	1.599 (1.289, 2.002)*	1.149 (0.881, 1.506)
Female	0.856 (0.765, 0.957)*	1.211 (1.017, 1.441)*	0.683 (0.558, 0.836)*
Smoker	1.393 (1.206, 1.610)*	1.629 (1.315, 2.022)	2.310 (1.805, 2.996)*
Body Mass Index	1.014 (1.008, 1.021)*	0.993 (0.983, 1.004)	1.008 (0.996, 1.021)
ABI < 0.9	1.348 (1.195, 1.521)*	1.068 (0.873, 1.312)	1.799 (1.453, 2.229)*
ACEI/ARB use	0.971 (0.868, 1.088)	1.081 (0.908, 1.291)	0.883 (0.727, 1.078)
Glucose (mg/dL)	1.000 (0.999, 1.001)	1.000 (0.998, 1.002)	0.999 (0.997, 1.001)
Hemoglobin A1c (%)	1.144 (1.096, 1.193)*	1.169 (1.095, 1.252)*	1.222 (1.125, 1.329)*
Hemoglobin (g/dL)	0.940 (0.910, 0.972)*	0.862 (0.814, 0.909)*	0.863 (0.809, 0.918)*
Calcium (mg/dL)	0.776 (0.701, 0.857)*	0.583 (0.495, 0.681)*	0.834 (0.690, 1.001)
Creatinine (mg/dL)	1.456 (1.313, 1.615)*	9.065 (7.192, 12.196)*	2.562 (2.055, 3.301)*
High BP	1.020 (0.882, 1.175)	1.297 (1.036, 1.612)*	0.946 (0.736, 1.218)
SBP (mmHg)	1.007 (1.003, 1.011)*	1.025 (1.019, 1.032)*	1.016 (1.009, 1.023)*
DPB (mmHg)	0.996 (0.915, 1.002)	0.996 (0.988, 1.004)	0.997 (0.988, 1.007)
Diabetes	-	1.535 (1.244, 1.903)*	1.330 (1.059, 1.671)*
Diabetes - ϕ_{11}	1.207 (1.046, 1.391)*	-	-
Diabetes - ϕ_{12}	1.006 (0.834, 1.213)	-	-
Diabetes - ϕ_{13}	0.836 (0.661, 1.055)	-	-
Hypertension		1.414 (1.056, 1.907)*	1.416 (1.043, 1.954)*
Hypertension - ϕ_{21}	1.639 (1.317, 2.041)*	-	-
Hypertension - ϕ_{22}	1.190 (0.864, 1.674)	-	-
Hypertension - ϕ_{23}	1.251 (0.845, 1.904)	-	-
CVD	-	1.575 (1.317, 1.894)*	2.632 (2.155, 3.248)*
CVD - ϕ_{31}	2.725 (2.405, 3.087)*	-	-
CVD - ϕ_{32}	1.751 (1.478, 2.083)*	-	-
CVD - ϕ_{33}	1.629 (1.314, 2.028)*	-	-
α_1	5.853 (4.018, 8.532)*	-	-
α_2	1.651 (1.004, 2.714)*	-	-
	Correlation/Covariance Estimate (95% CI)		
0.5pt σ_ν^2	2.238 (1.963, 2.589)*	-	-
$\zeta^{(1)}$	-	4.541 (3.285, 6.974)*	-
$\zeta^{(2)}$	-	-	6.958 (4.933, 11.052)*

1. * 95% credible interval (CI) does not include estimate of 0 or hazard ratio (HR) of 1;
2. * Reference group: Non-Hispanic White.

Table C.7: Simulation results from the proposed BM-JM and two simplified joint models, i.e., LT-JM (ignoring recurrent events) and RT-JM (ignoring longitudinal measurements) for sample size $n = 2000$ with censoring rates at about 30% and recurrent rate of about 70%. Given are bias, mean squared error (MSE), and average coverage probabilities (CP, in %) of the 95% credible intervals averaged over 200 datasets.

		BM-JM			LT-JM			RT-JM		
	True	Bias	MSE	CP	Bias	MSE	CP	Bias	MSE	CP
Longitudinal										
β_{1l}	0.60	−0.006	1.0×10^{-4}	94.5	−0.008	1.8×10^{-4}	92.0	-	-	-
ϕ_l	−1.30	0.019	2.1×10^{-4}	95.5	0.034	6.8×10^{-4}	92.5	-	-	-
γ	4.00	0.026	4.2×10^{-5}	96.5	−0.057	2.0×10^{-4}	90.0	-	-	-
Recurrent Events										
β_r	0.80	−0.034	0.002	93.0	-	-	-	−0.043	0.003	84.5
ϕ_1	1.20	−0.010	6.9×10^{-5}	96.5	-	-	-	−1.013	0.713	53.0
ϕ_2	−0.60	0.095	0.025	93.5	-	-	-	0.095	0.025	62.0
ϕ_3	−0.93	0.112	0.014	92.0	-	-	-	0.295	0.101	53.5
α_1	0.60	−0.035	0.003	95.5	-	-	-	−0.010	3.0×10^{-4}	99.0
α_2	−0.80	0.028	0.001	97.0	-	-	-	0.054	0.005	96.0
η_{r0}	0.80	−0.030	0.001	94.5	-	-	-	-	-	-
η_{r1}	0.60	−0.060	0.010	92.0	-	-	-	-	-	-
Competing-Risk Terminal Events										
$\beta_t^{(1)}$	−2.00	0.088	0.002	92.5	0.707	0.125	23.5	1.401	0.490	0
$\beta_t^{(2)}$	−1.50	0.093	0.004	92.0	0.806	0.289	22.0	1.127	0.565	0
$\phi_t^{(1)}$	−1.50	0.150	0.010	90.5	0.772	0.265	20.0	−1.599	1.136	0
$\phi_t^{(2)}$	−1.80	0.154	0.007	91.0	0.879	0.239	19.0	−1.771	0.969	0
$\eta_t^{(1)}$	2.00	−0.075	0.001	92.0	−0.742	0.138	19.5	-	-	-
$\eta_t^{(2)}$	1.70	−0.074	0.002	93.0	−0.862	0.257	17.0	-	-	-
$\zeta^{(1)}$	1.20	−0.012	1.0×10^{-4}	96.0	-	-	-	0.236	0.039	29.0
$\zeta^{(2)}$	1.50	0.011	5.4×10^{-5}	97.0	-	-	-	0.143	0.009	79.0
Variance/Correlation										
$\sigma^2_{b_0}$	1.25	−0.033	7.0×10^{-4}	94.5	−0.045	1.3×10^{-3}	93.5	-	-	-
$\sigma^2_{b_1}$	0.80	0.029	0.001	96.5	0.050	0.004	93.0	-	-	-
ρ_b	0.50	0.024	0.002	96.0	−0.082	0.027	88.5	-	-	-
σ^2_ε	1.32	0.004	9.2×10^{-6}	95.0	0.005	1.4×10^{-5}	93.0	-	-	-
σ^2_ν	1.44	−0.167	0.013	90.0	-	-	-	1.231	0.731	0

Table C.8: Simulation results from three marginal models: the longitudinal model (L-M), the recurrent model (R-T) and the competing-risk terminal model (T-M) for sample size $n = 2000$ with censoring rates at about 30% and recurrent rate of about 70%. Given are bias, mean squared error (MSE), and average coverage probabilities (CP, in %) of the 95% credible intervals averaged over 200 datasets.

		L-M			R-M			T-M		
	True	Bias	MSE	CP	Bias	MSE	CP	Bias	MSE	CP
Longitudinal										
β_{1l}	0.60	-0.013	4.8×10^{-4}	87.0	-	-	-	-	-	-
ϕ_l	-1.30	0.066	0.003	76.0	-	-	-	-	-	-
γ	4.00	-0.286	0.005	0	-	-	-	-	-	-
Recurrent Events										
β_r	0.80	-	-	-	0.063	0.006	73.0	-	-	-
ϕ_1	1.20	-	-	-	0.248	0.043	37.0	-	-	-
ϕ_2	-0.60	-	-	-	0.608	1.027	0	-	-	-
ϕ_3	-0.93	-	-	-	0.709	0.582	0	-	-	-
α_1	0.60	-	-	-	-0.286	0.227	9.5	-	-	-
α_2	-0.80	-	-	-	-0.091	0.013	94.5	-	-	-
Competing-Risk Terminal Events										
$\beta_t^{(1)}$	-2.00	-	-	-	-	-	-	1.630	0.664	0
$\beta_t^{(2)}$	-1.50	-	-	-	-	-	-	1.362	0.824	0
$\phi_t^{(1)}$	-1.50	-	-	-	-	-	-	-0.024	2.6×10^{-4}	57.0
$\phi_t^{(2)}$	-1.80	-	-	-	-	-	-	-0.401	0.050	73.5
Variance/Correlation										
$\sigma^2_{b_0}$	1.25	-0.037	8.9×10^{-4}	86.0	-	-	-	-	-	-
$\sigma^2_{b_1}$	0.80	-0.142	0.032	59.5	-	-	-	-	-	-
ρ_b	0.50	-0.102	0.042	47.5	-	-	-	-	-	-
σ^2_ε	1.32	0.007	3.0×10^{-5}	92.0	-	-	-	-	-	-
σ^2_ν	1.44	-	-	-	1.087	0.570	0	-0.250	0.030	66.5

Table C.9: Time-dynamic AUC and BS values from the proposed BM-JM and two simplified joint models, i.e., LT-JM (ignoring recurrent events) and RT-JM (ignoring longitudinal measurements) for the test sample of $n_p = 50$ subjects. The results are based on 200 Monte Carlo runs. Note that $\pi^{(w)}$ denotes the cumulative incidence probability calculated for the w-th competing risk where $w = 1, 2$.

		$\text{AUC}_{\pi^{(1)}}$			$\text{AUC}_{\pi^{(2)}}$			$\text{BS}_{\pi^{(1)}}$			$\text{BS}_{\pi^{(2)}}$		
t	Δt	BM-JM	LT-JM	RT-JM	BM-JM	LT-JM	RT-JM	BM-JM	LT-JM	RT-JM	BM-JM	LT-JM	RT-JM
0	0.2	0.931	0.892	0.901	0.943	0.910	0.931	0.036	0.039	0.038	0.051	0.061	0.053
	0.4	0.886	0.856	0.845	0.905	0.849	0.867	0.076	0.084	0.078	0.093	0.111	0.108
	0.6	0.832	0.810	0.827	0.867	0.804	0.856	0.121	0.132	0.148	0.136	0.157	0.145
0.3	0.2	0.900	0.852	0.892	0.900	0.785	0.873	0.062	0.077	0.075	0.081	0.091	0.091
	0.4	0.840	0.804	0.811	0.868	0.757	0.841	0.119	0.128	0.171	0.126	0.150	0.143
	0.6	0.785	0.755	0.673	0.812	0.726	0.715	0.164	0.186	0.223	0.167	0.198	0.196
0.6	0.2	0.870	0.813	0.757	0.879	0.765	0.758	0.097	0.128	0.277	0.096	0.131	0.261
	0.4	0.770	0.676	0.500	0.744	0.658	0.500	0.189	0.244	0.295	0.197	0.259	0.341

Bibliography

Albert, P. S. and Shih, J. H. (2010). An approach for jointly modeling multivariate longitudinal measurements and discrete time-to-event data. *The Annals of Applied Statistics*, 4(3):1517.

Baer, D. R., Lawson, A. B., and Joseph, J. E. (2021). Joint space–time bayesian disease mapping via quantification of disease risk association. *Statistical Methods in Medical Research*, 30(1):35–61.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2015). *Hierarchical modeling and analysis for spatial data (2nd Edition)*. CRC Press.

Bello, A. K., Okpechi, I. G., Osman, M. A., Cho, Y., Htay, H., Jha, V., Wainstein, M., and Johnson, D. W. (2022). Epidemiology of haemodialysis outcomes. *Nature Reviews Nephrology*, 18(6):378–395.

Blanche, P., Dartigues, J. F., and Jacqmin-Gadda, H. (2013). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in Medicine*, 32(30):5381–5397.

Blanche, P., Proust-Lima, C., Loubère, L., Berr, C., Dartigues, J. F., and Jacqmin-Gadda, H. (2015). Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*, 71(1):102–113.

Boulieri, A., Liverani, S., de Hoogh, K., and Blangiardo, M. (2017). A space–time multivariate bayesian model to analyse road traffic accidents by severity. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(1):119–139.

Cai, J., Fan, J., Zhou, H., and Zhou, Y. (2007). Hazard models with varying coefficients for multivariate failure time data. *Annals of Statistics*, 35(1):324–354.

Campos, E., Scheffler, A., Telesca, D., Sugar, C., DiStefano, C., Jeste, S., Levin, A. R., Naples, A., Webb, S. J., Shic, F., et al. (2022). Multilevel hybrid principal components analysis for region-referenced functional electroencephalography data. *Statistics in Medicine*, 41(19):3737–3757.

Cheng, W., Gill, G. S., Dasu, R., Xie, M., Jia, X., and Zhou, J. (2017). Comparison of multivariate poisson lognormal spatial and temporal crash models to identify hot spots of intersections based on crash types. *Accident Analysis & Prevention*, 99:330–341.

Chiou, J., Chen, Y., and Yang, Y. (2014). Multivariate functional principal component analysis: A normalization approach. *Statistica Sinica*, 24(4):1571–1596.

Congdon, P. (2004). A multivariate model for spatio-temporal health outcomes with an application to suicide mortality. *Geographical Analysis*, 36(3):234–258.

Coresh, J., Selvin, E., Stevens, L. A., Manzi, J., Kusek, J. W., Eggers, P., van Lente, F., and Levey, A. S. (2007). Prevalence of chronic kidney disease in the united states. *JAMA*, 298(17):2038–2047.

Cox, D. D. (1993). An analysis of bayesian inference for nonparametric regression. *The Annals of Statistics*, 21(2):903–923.

Crainiceanu, C., Ruppert, D., and Wand, M. P. (2005). Bayesian analysis for penalized spline regression using winbugs. *Journal of Statistical Software*, 14(14):1–24.

Crainiceanu, C. M., Ruppert, D., Carroll, R. J., Joshi, A., and Goodner, B. (2007). Spatially adaptive bayesian penalized splines with heteroscedastic errors. *Journal of Computational and Graphical Statistics*, 16(2):265–288.

Crainiceanu, C. M., Staicu, A. M., and Di, C. (2009). Generalized multilevel functional regression. *Journal of the American Statistical Association*, 104(488):1550–1561.

de Jager, D., Grootendorst, D., Jager, K., van Dijk, P., Tomas, L., Ansell, D., Collart, F., Finne, P., Heaf, J., De Meester, J., Wetzels, J., Rosendaal, F., and Dekker, F. (2009).

Cardiovascular and noncardiovascular mortality among patients starting dialysis. *Journal of the American Medical Association*, 302(16):1782–1789.

Di Salvo, F., Ruggieri, M., and Plaia, A. (2015). Functional principal component analysis for multivariate multidimensional environmental data. *Environmental and Ecological Statistics*, 22(4):739–757.

Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science*, 11(2):89–121.

Erickson, K. F., Zhao, B., Niu, J., Winkelmayer, W. C., Bhattacharya, J., Chertow, G. M., and Ho, V. (2019). Association of hospitalization and mortality among patients initiating dialysis with hemodialysis facility ownership and acquisitions. *JAMA Network Open*, 2(5):e193987–e193987.

Estes, J. P., Nguyen, D. V., Chen, Y., Dalrymple, L. S., Rhee, C. M., Kalantar-Zadeh, K., and Şentürk, D. (2018). Time-dynamic profiling with application to hospital readmission among patients on dialysis. *Biometrics*, 74(4):1383–1394.

Estes, J. P., Nguyen, D. V., Dalrymple, L. S., Mu, Y., and Şentürk, D. (2016). Time-varying effect modeling with longitudinal data truncated by death: conditional models, interpretations, and inference. *Statistics in Medicine*, 35(11):1834–1847.

Feldman, H. I., Appel, L. J., Chertow, G. M., Cifelli, D., Cizman, B., Daugirdas, J., Fink, J. C., Franklin-Becker, E. D., Go, A. S., Hamm, L. L., et al. (2003). The chronic renal insufficiency cohort (cric) study: design and methods. *Journal of the American Society of Nephrology*, 14(S2):S148–S153.

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509.

Foley, R. N., Chen, S., Solid, C. A., Gilbertson, D. T., and Collins, A. J. (2014). Early mortality in patients starting dialysis appears to go unregistered. *Kidney International*, 86(2):392–398.

Gelfand, A. E., Banerjee, S., and Gamerman, D. (2005). Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics: The Official Journal of the International Environmetrics Society*, 16(5):465–479.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Gerds, T. A. and Schumacher, M. (2006). Consistent estimation of the expected brier score in general survival models with right-censored event times. *Biometrical Journal*, 48(6):1029–1040.

Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E., and Hsu, C. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *New England Journal of Medicine*, 351(13):1296–1305.

Gould, A. L., Boye, M. E., Crowther, M. J., Ibrahim, J. G., Quartey, G., Micallef, S., and Bois, F. Y. (2015). Joint modeling of survival and longitudinal non-survival data: current methods and issues. *Statistics in Medicine*, 34(14):2181–2195.

Grams, M. E., Chow, E. K., Segev, D. L., and Coresh, J. (2013). Lifetime incidence of ckd stages 3-5 in the united states. *American Journal of Kidney Diseases*, 62(2):245–252.

Gueorguieva, R., Rosenheck, R., and Lin, H. (2012). Joint modelling of longitudinal outcome and interval-censored competing risk dropout in a schizophrenia clinical trial. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 175(2):417–433.

Guo, Y., Sun, D., and Sun, J. (2022). Inference of a time-varying coefficient regression model for multivariate panel count data. *Journal of Multivariate Analysis*, 192:105047.

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113(522):649–659.

He, K., Lian, H., Ma, S., and Huang, J. Z. (2018). Dimensionality reduction and variable selection in multivariate varying-coefficient models with a large number of covariates. *Journal of the American Statistical Association*, 113(522):746–754.

Hepler, S. A., Waller, L. A., and Kline, D. M. (2021). A multivariate spatiotemporal change-point model of opioid overdose deaths in ohio. *The Annals of Applied Statistics*, 15(3):1329–1342.

Hickson, L. J., Thorsteinsdottir, B., Ramar, P., Reinalda, M. S., Crowson, C. S., Williams, A. W., Albright, R. C., Onuigbo, M. A., Rule, A. D., and Shah, N. D. (2018). Hospital readmission among new dialysis patients associated with young age and poor functional status. *Nephron*, 139(1):1–12.

Hillege, H. L., Nitsch, D., Pfeffer, M. A., Swedberg, K., McMurray, J. J., Yusuf, S., Granger, C. B., Michelson, E. L., Ostergren, J., Cornel, J. H., et al. (2006). Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. *Circulation*, 113(5):671–678.

Jacques, J. and Preda, C. (2014). Model-based clustering for multivariate functional data. *Computational Statistics & Data Analysis*, 71:92–106.

Jiang, J., Lin, H., Zhong, Q., and Li, Y. (2022). Analysis of multivariate non-gaussian functional data: A semiparametric latent process approach. *Journal of Multivariate Analysis*, 189:104888.

Jin, X., Banerjee, S., and Carlin, B. P. (2007). Order-free co-regionalized areal data models with application to multiple-disease mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(5):817–838.

Jullion, A. and Lambert, P. (2007). Robust specification of the roughness penalty prior distribution in spatially adaptive bayesian p-splines models. *Computational Statistics & Data Analysis*, 51(5):2542–2558.

Kahaner, D., Moler, C., and Nash, S. (1989). *Numerical methods and software*. Prentice-Hall Inc.

Kind, A. J. and Buckingham, W. R. (2018). Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. *The New England Journal of Medicine*, 378(26):2456.

Krivobokova, T., Kneib, T., and Claeskens, G. (2010). Simultaneous confidence bands for penalized spline estimators. *Journal of the American Statistical Association*, 105(490):852–863.

Król, A., Ferrer, L., Pignon, J. P., Proust-Lima, C., Ducreux, M., Bouché, O., Michiels, S., and Rondeau, V. (2016). Joint model for left-censored longitudinal data, recurrent events and terminal event: Predictive abilities of tumor burden for cancer evolution with application to the ffcd 2000–05 trial. *Biometrics*, 72(3):907–916.

Kürüm, E., Kwan, B., Qian, Q., Banerjee, S., Rhee, C. M., Nguyen, D. V., and Şentürk, D. (2024). A bayesian joint model of longitudinal kidney disease progression, recurrent cardiovascular events, and terminal event in patients with chronic kidney disease. *Statistics in Biosciences*, 1(1):1–27.

Kürüm, E., Li, R., Shiffman, S., and Yao, W. (2016). Time-varying coefficient models for joint modeling binary and continuous outcomes in longitudinal data. *Statistica Sinica*, 26(3):979.

Lang, S. and Brezger, A. (2004). Bayesian p-splines. *Journal of Computational & Graphical Statistics*, 13(1):183–212.

Li, Y., Nguyen, D. V., Banerjee, S., Rhee, C. M., Kalantar-Zadeh, K., Kürüm, E., and Şentürk, D. (2021). Multilevel modeling of spatially nested functional data: Spatiotemporal patterns of hospitalization rates in the us dialysis population. *Statistics in Medicine*, 40(17):3937–3952.

Li, Y., Nguyen, D. V., Chen, Y., Rhee, C. M., Kalantar-Zadeh, K., and Şentürk, D. (2018).

Modeling time-varying effects of multilevel risk factors of hospitalizations in patients on dialysis. *Statistics in Medicine*, 37(30):4707–4720.

Li, Y., Nguyen, D. V., Kürüm, E., Rhee, C. M., Banerjee, S., and Şentürk, D. (2022). Multilevel varying coefficient spatiotemporal model. *Stat*, 11(1):e438.

Liu, L. and Huang, X. (2009). Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 58(1):65–81.

Matsushita, K., van der Velde, M., Astor, B. C., Woodward, M., Levey, A. S., de Jong, P. E., Coresh, J., and Gansevoort, R. T. (2010). Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet*, 375(9731):2073–2081.

Muntner, P., He, J., Hamm, L., Loria, C., and Whelton, P. K. (2002). Renal insufficiency and subsequent death resulting from cardiovascular disease in the united states. *Journal of the American Society of Nephrology*, 13(3):745–753.

Noordzij, M. and Jager, K. J. (2014). Increased mortality early after dialysis initiation: a universal phenomenon. *Kidney International*, 85(1):12–14.

Plummer, M. (2017). Jags version 4.3. 0 user manual [computer software manual]. sourceforge net/projects/mcmc-jags/files. *Manuals/4 x*, 2.

Qian, Q., Nguyen, D. V., Telesca, D., Kurum, E., Rhee, C. M., Banerjee, S., Li, Y., and Senturk, D. (2024). Multivariate spatiotemporal functional principal component analysis for modeling hospitalization and mortality rates in the dialysis population. *Biostatistics*, 25(3):718–735.

Quick, H., Banerjee, S., and Carlin, B. P. (2013). Modeling temporal gradients in regionally aggregated california asthma hospitalization data. *The Annals of Applied Statistics*, 7(1):154.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer.

Ren, X., Wang, J., and Luo, S. (2021). Dynamic prediction using joint models of longitudinal and recurrent event data: a bayesian perspective. *Biostatistics & Epidemiology*, 5(2):250–266.

Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, 67(3):819–829.

Robins, J. M. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology: Methodological Issues*, 1(1):297–331.

Ruggieri, M., Plaia, A., and Di Salvo, F. (2018). Comparing spatial and spatio-temporal fpca to impute large continuous gaps in space. *Classification, (Big) Data Analysis & Statistical Learning*, 1(1):201–208.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational & Graphical Statistics*, 11(4):735–757.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric regression*. Cambridge University Press.

Scheffler, A., Telesca, D., Li, Q., Sugar, C. A., Distefano, C., Jeste, S., and Şentürk, D. (2020). Hybrid principal components analysis for region-referenced longitudinal functional eeg data. *Biostatistics*, 21(1):139–157.

Schoop, R., Beyersmann, J., Schumacher, M., and Binder, H. (2011). Quantifying the predictive accuracy of time-to-event models in the presence of competing risks. *Biometrical Journal*, 53(1):88–112.

Schoop, R., Graf, E., and Schumacher, M. (2008). Quantifying the predictive performance of prognostic models for censored survival data with time-dependent covariates. *Biometrics*, 64(2):603–610.

Skupien, J., Warram, J. H., Smiles, A. M., Stanton, R. C., and Krolewski, A. S. (2016). Patterns of estimated glomerular filtration rate decline leading to end-stage renal disease in type 1 diabetes. *Diabetes Care*, 39(12):2262–2269.

Staicu, A. M., Crainiceanu, C. M., and Carroll, R. J. (2010). Fast methods for spatially correlated multilevel functional data. *Biostatistics*, 11(2):177–194.

Thiébaut, R., Jacqmin-Gadda, H., Babiker, A., Commenges, D., and Collaboration, C. (2005). Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of cd4+ cell count and hiv rna viral load in response to treatment of hiv infection. *Statistics in Medicine*, 24(1):65–82.

Uno, H., Cai, T., Tian, L., and Wei, L. J. (2007). Evaluating prediction rules for t-year survivors with censored regression models. *Journal of the American Statistical Association*, 102(478):527–537.

USRDS (2021). United States Renal Data System 2021 Annual Data Report: "Epidemiology of Kidney Disease in the United States". Technical report, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD.

USRDS (2022). United States Renal Data System 2022 Annual Data Report: "Epidemiology of Kidney Disease in the United States". Technical report, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD.

Wild, C. and Yee, T. (1996). Additive extensions to generalized estimating equation methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(4):711–725.

Williamson, P. R., Kolamunnage-Dona, R., Philipson, P., and Marson, A. G. (2008). Joint modelling of longitudinal and competing risks data. *Statistics in Medicine*, 27(30):6426–6438.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics*, 53(1):330–339.

Yang, W., Xie, D., Anderson, A. H., Joffe, M. M., Greene, T., Teal, V., Hsu, C., Fink, J. C., He, J., Lash, J. P., et al. (2014). Association of kidney disease outcomes with risk factors for ckd: findings from the chronic renal insufficiency cohort (cric) study. *American Journal of Kidney Diseases*, 63(2):236–243.

Yee, T. W. and Mackenzie, M. (2002). Vector generalized additive models in plant ecology. *Ecological Modelling*, 157(2-3):141–156.

Yee, T. W. and Wild, C. (1996). Vector generalized additive models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(3):481–493.

Zhang, F., Li, R., Lian, H., and Bandyopadhyay, D. (2021). Sparse reduced-rank regression for multivariate varying-coefficient models. *Journal of Statistical Computation & Simulation*, 91(4):752–767.

Zhang, S., Sun, D., He, C. Z., and Schootman, M. (2006). A bayesian semi-parametric model for colorectal cancer incidences. *Statistics in Medicine*, 25(2):285–309.

Zheng, Y., Cai, T., Jin, Y., and Feng, Z. (2012). Evaluating prognostic accuracy of biomarkers under competing risk. *Biometrics*, 68(2):388–396.

Zhu, H., Li, R., and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *Annals of Statistics*, 40(5):2634.