

Supplementary Materials for “Bayesian multivariate joint modeling of longitudinal, recurrent, and competing risk terminal events in patients with chronic kidney disease”

QI QIAN, DANH V. NGUYEN, ESRA KURUM, SUDIPTO BANERJEE,
CONNIE M. RHEE, and DAMLA SENTURK

Appendix A. Prior Distributions and MCMC Implementation

For both data analysis and simulation studies, normal priors with mean zero and variance σ_\star^2 were utilized in model fits for the fixed-effects parameters $\beta = (\beta_\ell^\top, \beta_r^\top, \beta_t^{(1)\top}, \beta_t^{(2)\top})$, $\phi = (\phi_\ell^\top, \phi_r^\top, \phi_t^{(1)\top}, \phi_t^{(2)\top})$, γ , and $\alpha = (\alpha_1, \alpha_2)$, and the association parameters $\eta = (\eta_{r0}, \eta_{r1}, \eta_t^{(1)}, \eta_t^{(2)})$ and $\zeta = (\zeta^{(1)}, \zeta^{(2)})$, where the hyperparameters were drawn from $\sigma_\star^2 \sim \text{IG}(1, 0.005)$. For the variance terms, $\sigma^2 = (\sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_\varepsilon^2, \sigma_\nu^2)$, IG priors with parameters ($a_{1\star} = 0.001, a_{2\star} = 0.001$) were utilized (\star 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. We note that with respect to the number of knots used for the baseline hazard, we followed the results and recommendations provided by Eilers and Marx, Lang and Brezger, and Eilers and Marx [31, 35, 36] 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 [37] 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 supplementary Figures S1-S8. 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.

Appendix B. 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 (6) of the main paper, 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 S1, 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 h th iteration, for $h = 1, \dots, H$ (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 (6) (of the main paper) is of nonstandard form, and thus a more sophisticated approach is required to sample from it. Note that

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

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 $\tilde{\mathbf{u}}_p$ from a multivariate t distribution with four degrees of freedom (MVT₄) centered at the empirical Bayesian estimates $\hat{\mathbf{u}}_p = \text{argmax}_{\mathbf{u}} \log p(\mathbf{u}_p \mid \cup_{w=1}^2 (T_p^{*(w)} > t), \mathcal{Y}_p(t), \mathcal{R}_p(t); \hat{\boldsymbol{\theta}})$, with scale matrix

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

where \mathcal{I} denotes Fisher information matrix.

The rationale for employing a MVT₄ proposal is two-fold. First, as more longitudinal measurements and recurrent event information is incorporated, the survival and random effect components in (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 $\hat{\mathbf{u}}_p$ and $\tilde{\mathbf{u}}_p$ are obtained, we

proceed to calculate the Metropolis-Hastings acceptance ratio

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

If a randomly generated $\tilde{\Lambda} \sim \text{Unif}(0, 1)$ is less than Λ , i.e., $\tilde{\Lambda} < \Lambda$, then proposal $\tilde{\mathbf{u}}_p$ is accepted $\mathbf{u}_p^{(h)} \equiv \tilde{\mathbf{u}}_p$, otherwise, $\mathbf{u}_p^{(h)} \equiv \hat{\mathbf{u}}_p$.

Step 3: Combining (4) and (6) (of the main paper), we denote the cumulative incidence probability for the w th 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

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

where $\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ and $S_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ are calculated using (7) of the main paper 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 $S_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})$ 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 $\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}), h = 1, \dots, H\}$. Moreover, credible intervals can be obtained using the Monte Carlo sample percentiles.

Appendix C. 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

(ROC) curves (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 $\{(T_p, \delta_p, \pi_p^{(w)}(t, s)), p = 1, \dots, n_p\}$, 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 w th 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 ($p, p' = 1, \dots, n_p$ 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 w th competing risk $BS^{(w)}(t, s) = \mathbb{E} \left[\left\{ D^{(w)}(t, s) - \pi^{(w)}(t, s) \right\}^2 \mid T^* > t \right]$, 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,

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

and

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

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 $\tilde{D}_p^{(w)}(t, s)$ to account for the indicator obtained only from those subjects whose outcomes are observed and known. Specifically, $\tilde{D}_p^{(w)}(t, s) = I_{(t < T_p \leq s, \delta_p = w)}$ 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 $\widehat{W}_p(t, s) = I_{(T_p > s)} / \widehat{G}(s | t) + \{I_{(t < T_p \leq s)} I_{(T_p^* \leq C_p)}\} / \widehat{G}(T_p | t)$, where $\widehat{G}(s | t)$ and $\widehat{G}(T_p | t)$ 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 censored. 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).

Appendix D. Details on Comparative Models

We compared the proposed BM-JM with other well-established models that overlook the joint outcomes structure. As mentioned in Section 4.3 of the main paper, 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):

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

Recurrent model (R-M):

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

Competing-risk terminal model (T-M):

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

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) :

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

$$h_i^{(1)}(t) = h_{i0}^{(1)}(t) \exp \left(\mathbf{X}_i^\top \boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^\top \boldsymbol{\phi}_t^{(1)} + \eta_t^{(1)} \xi_i(t) \right),$$

$$h_i^{(2)}(t) = h_{i0}^{(2)}(t) \exp \left(\mathbf{X}_i^\top \boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^\top \boldsymbol{\phi}_t^{(2)} + \eta_t^{(2)} \xi_i(t) \right).$$

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

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

$$h_i^{(1)}(t) = h_{i0}^{(1)}(t) \exp \left(\mathbf{X}_i^\top \boldsymbol{\beta}_t^{(1)} + \mathbf{Z}_i^\top \boldsymbol{\phi}_t^{(1)} + \zeta^{(1)} \nu_i \right),$$

$$h_i^{(2)}(t) = h_{i0}^{(2)}(t) \exp \left(\mathbf{X}_i^\top \boldsymbol{\beta}_t^{(2)} + \mathbf{Z}_i^\top \boldsymbol{\phi}_t^{(2)} + \zeta^{(2)} \nu_i \right).$$

Note that in the LT-JM and RT-JM, the longitudinal and competing-risk terminal submodels 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.

Appendix E. 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 D. 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 $\mathbf{b}_i = (b_{i0}, b_{i1})^\top \sim N(\mathbf{0}, \boldsymbol{\Sigma}_b)$ with $\boldsymbol{\Sigma}_b = [\sigma_{b0}^2, \sigma_{01}; \sigma_{01}, \sigma_{b1}^2]$, $\sigma_{01} = \rho_b \sigma_{b0} \sigma_{b1}$, $\rho_b = 0.50$, $\sigma_{b0}^2 = 1.25$, and $\sigma_{b1}^2 = 0.80$, and the random error $\varepsilon_i(t) \sim N(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 = 1.32$.
- *Recurrent processes:* $r_{ij}(g) = h_{r0}(g) \exp(X_i \beta_r + Z_i \phi_{rj} + \sum_{m=0}^{j-1} \alpha_m + \eta_{r1} b_{i0} + \eta_{r2} b_{i1} + \nu_i)$, 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 $\nu_i \sim N(0, \sigma_\nu^2)$ with $\sigma_\nu^2 = 1.44$.

- *Competing-risk terminal processes:* $h_i^{(w)}(t) = h_{t_0}^{(w)}(t) \exp(X_i \beta_t^{(w)} + Z_i \phi_t^{(w)} + \eta_t^{(w)} \xi_i(t) + \zeta^{(w)} \nu_i)$ ($w = 1, 2$), 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 $h_{t_0}^{(w)}(t) = 1.50t^{0.50}$ ($w = 1, 2$), respectively. This results in a probability of having one, two or three recurrent events of approximately .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 $\widehat{\pi}_p^{(w)}(t, s) = \text{median}\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}), h = 1, \dots, H\}$ over $H = 200$ Monte Carlo samples using the proposed Monte Carlo simulation procedure detailed in Appendix B and Table S1. 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 3.2 of the main paper and Appendix C, and the results are summarized in Table S9.

References

- Blanche P, Dartigues JF, 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 JF, 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.
- Gelman A, Rubin DB (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4): 457–472.
- Gerds TA, 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.

- Kahaner D, Moler C, Nash S (1989). *Numerical Methods and Software*. Prentice-Hall, Inc.
- Kürüm E, Kwan B, Qian Q, Banerjee S, Rhee CM, Nguyen DV, et al. (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*, In print.
- Plummer M (2017). Jags version 4.3. 0 user manual [computer software manual]. sourceforge net/projects/mcmc-jags/files. *Manuals/4 x, 2*.
- Rizopoulos D (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, 67(3): 819–829.
- Robins JM, Rotnitzky A (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In: *AIDS epidemiology: methodological issues*, 297–331. Springer.
- Schoop R, Beyersmann J, Schumacher M, 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, Schumacher M (2008). Quantifying the predictive performance of prognostic models for censored survival data with time-dependent covariates. *Biometrics*, 64(2): 603–610.
- Uno H, Cai T, Tian L, Wei LJ (2007). Evaluating prediction rules for t-year survivors with censored regression models. *Journal of the American Statistical Association*, 102(478): 527–537.
- Zheng Y, Cai T, Jin Y, Feng Z (2012). Evaluating prognostic accuracy of biomarkers under competing risk. *Biometrics*, 68(2): 388–396.

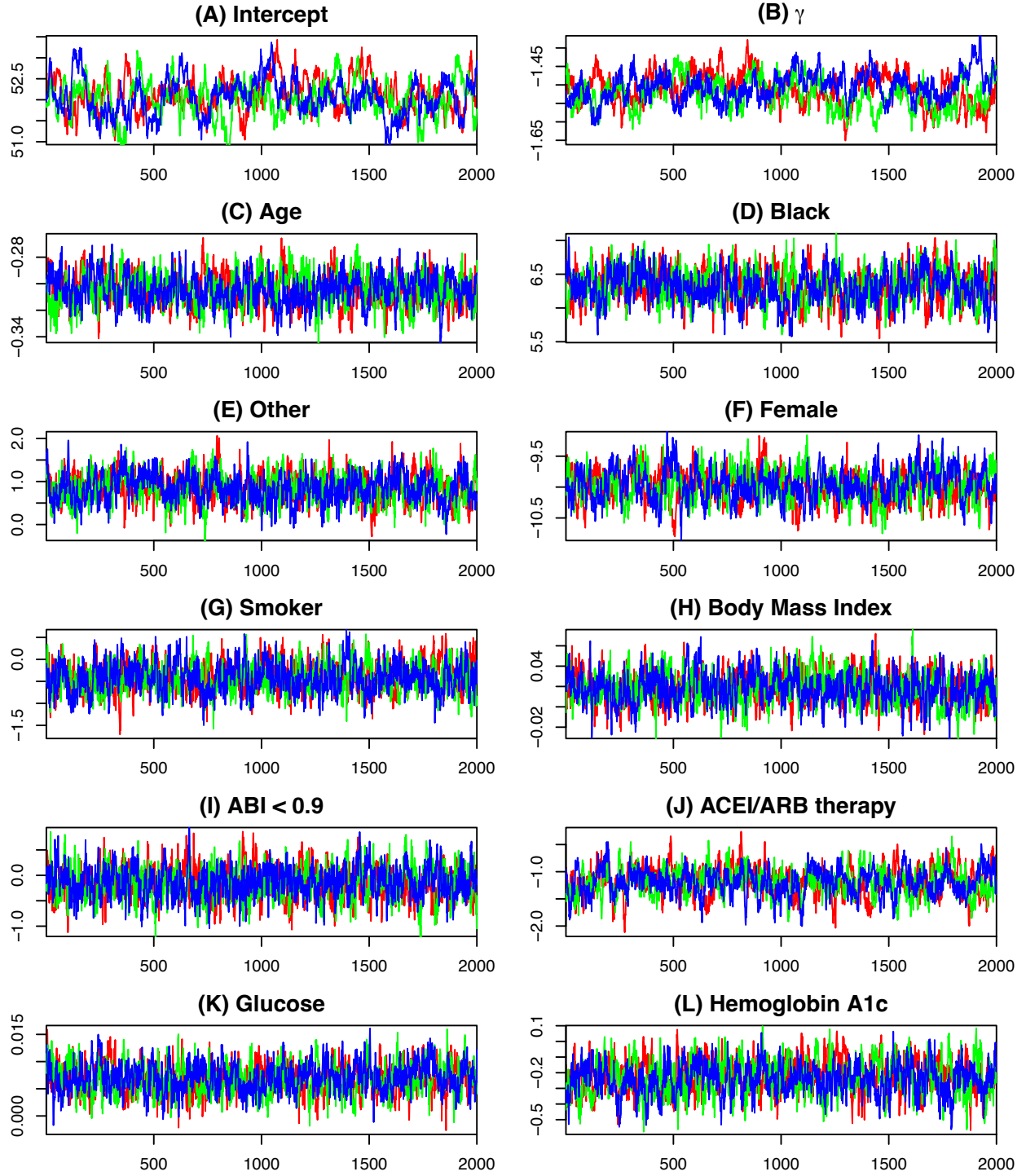


Figure S1: Trace plots for longitudinal submodel parameters (part 1).

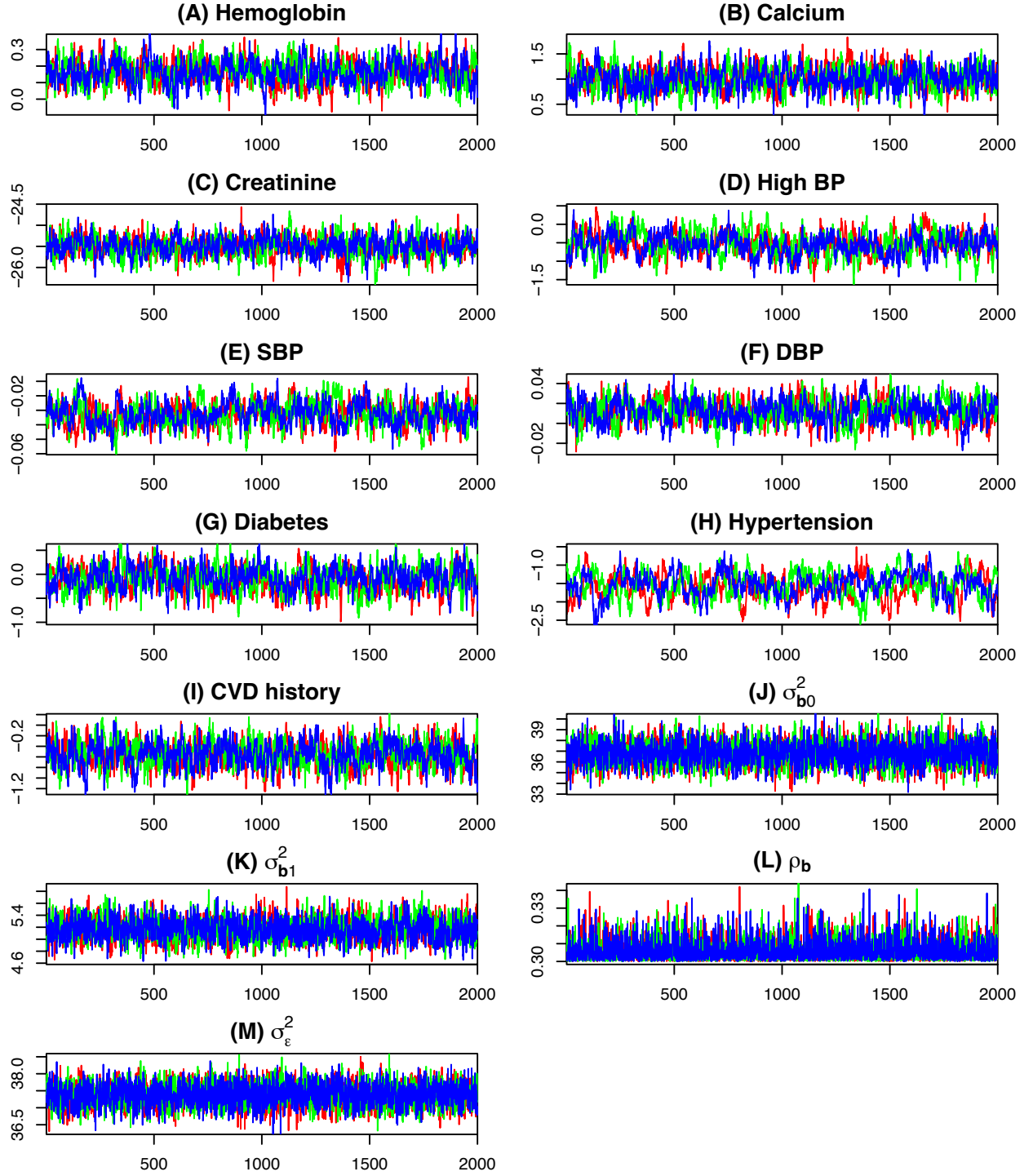


Figure S2: Trace plots for longitudinal submodel parameters (part 2).

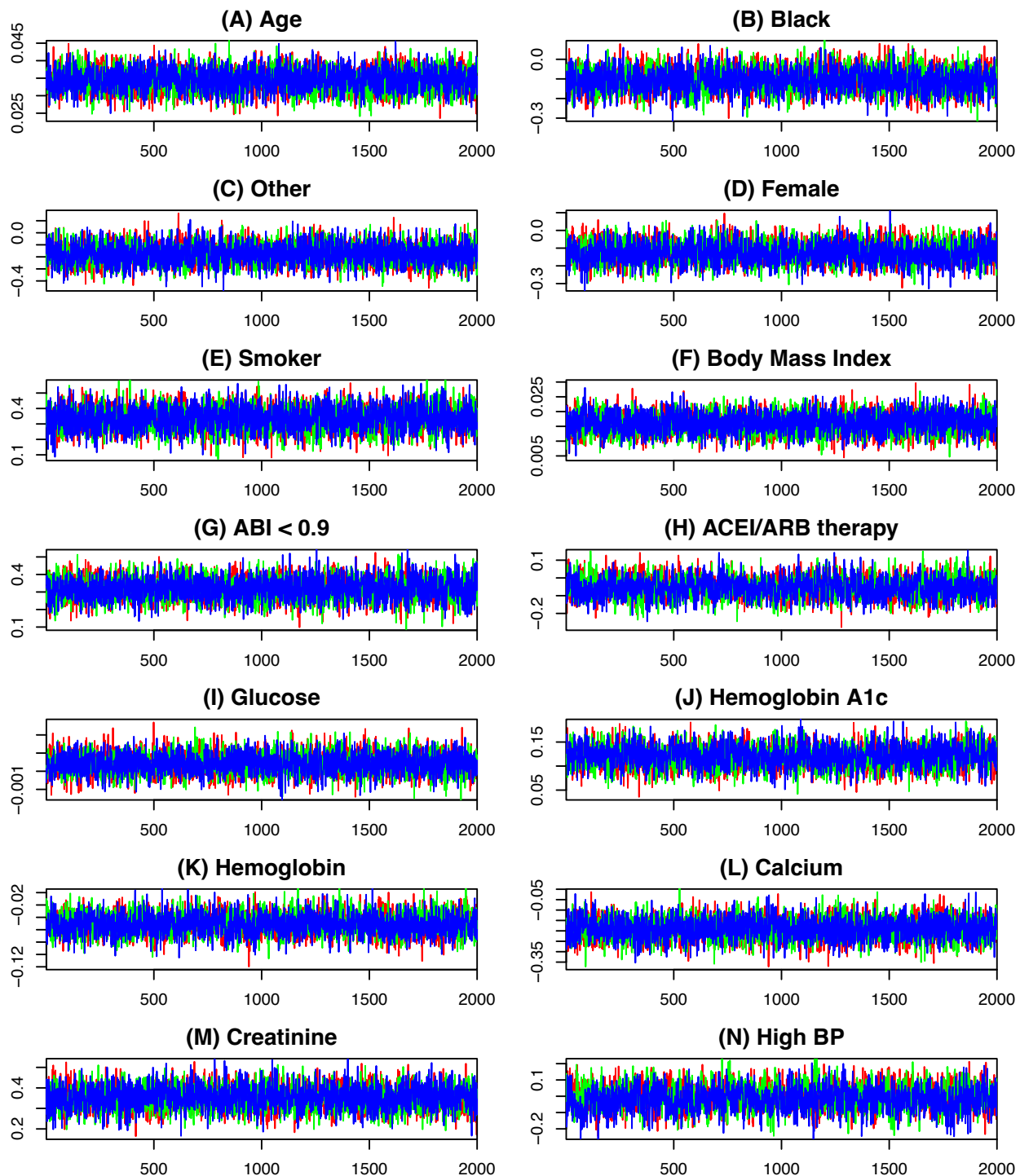


Figure S3: Trace plots for recurrent events submodel parameters (part 1).

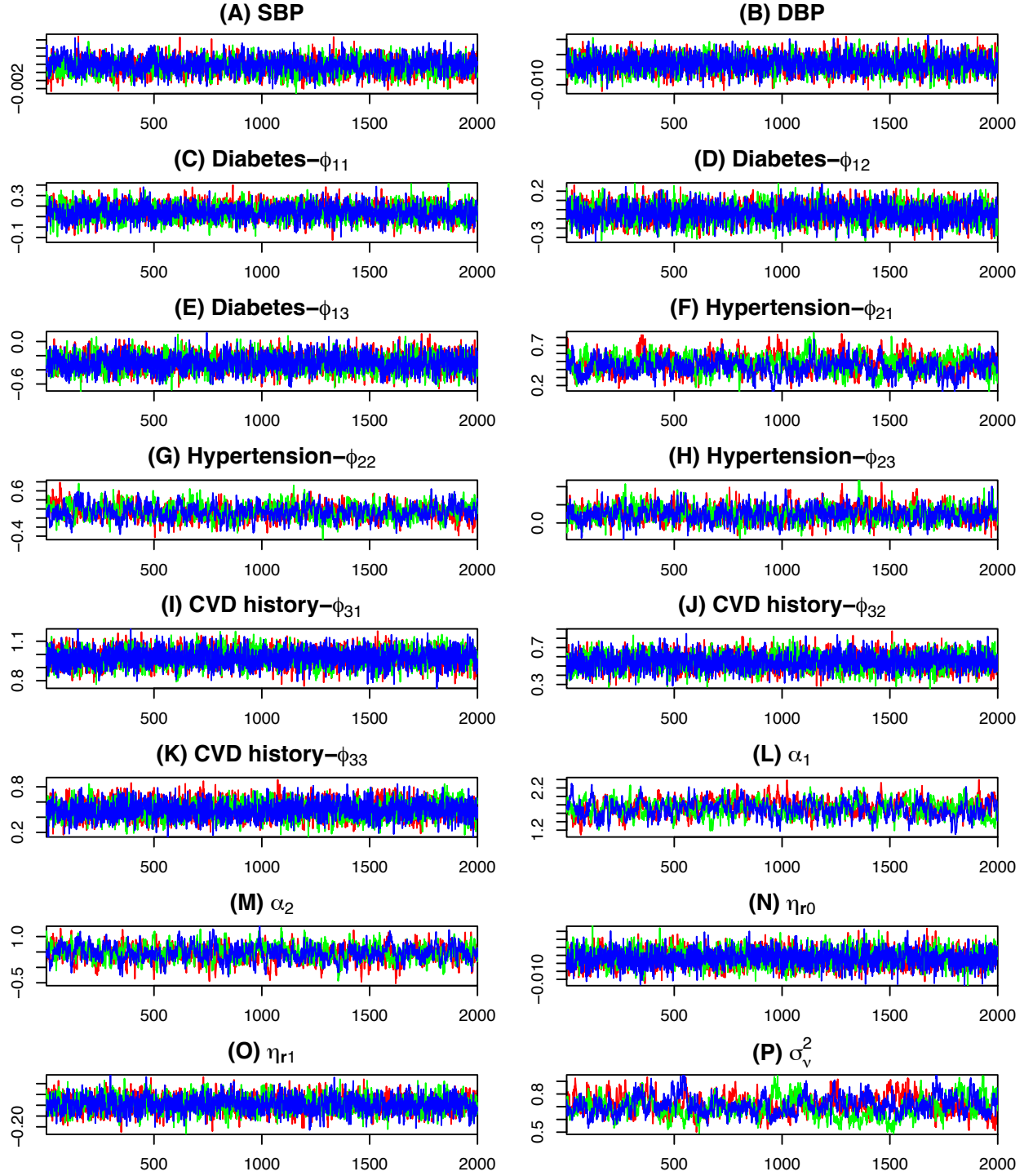


Figure S4: Trace plots for recurrent events submodel parameters (part 2).

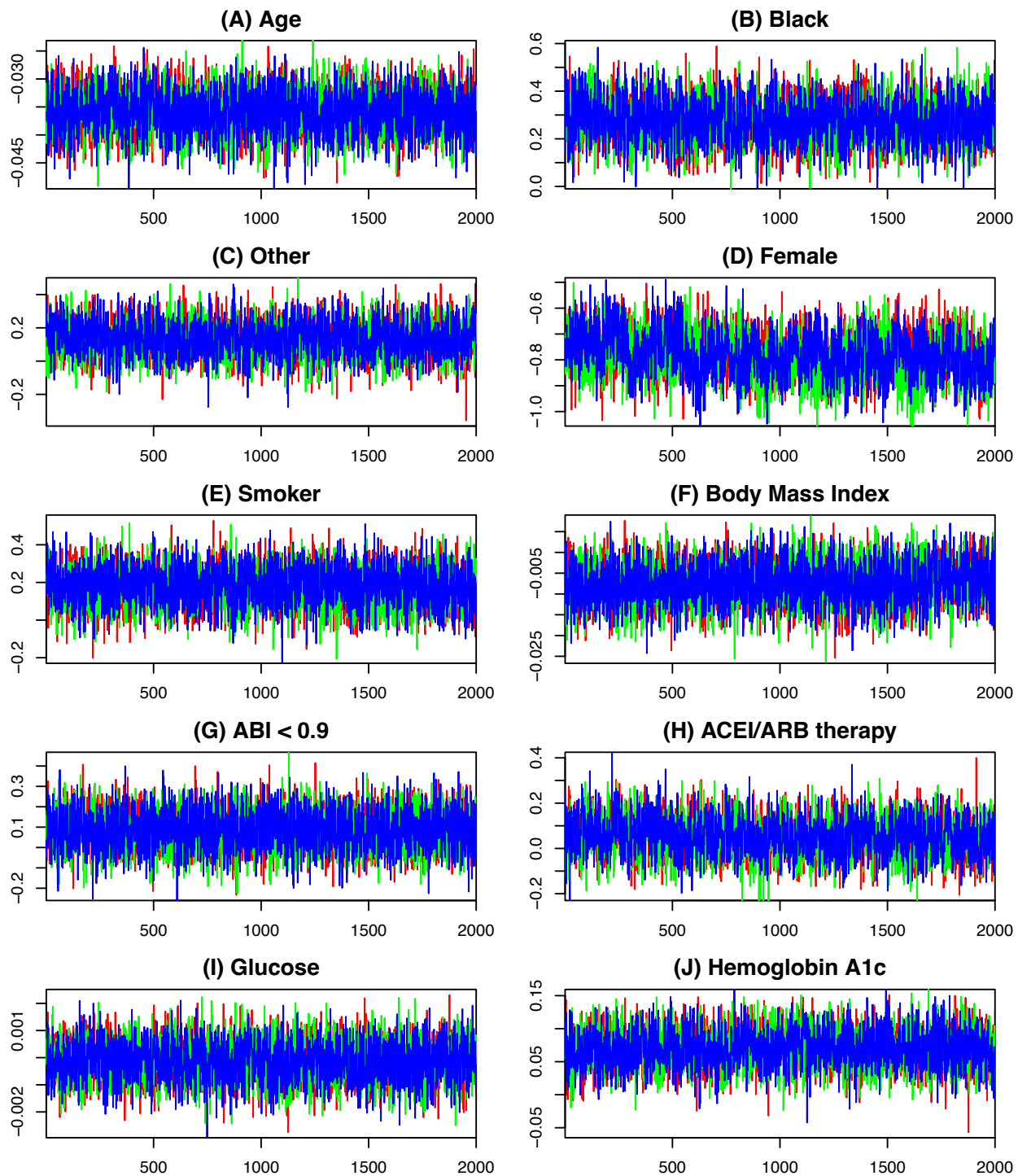


Figure S5: Trace plots for ESKD submodel parameters (part 1).

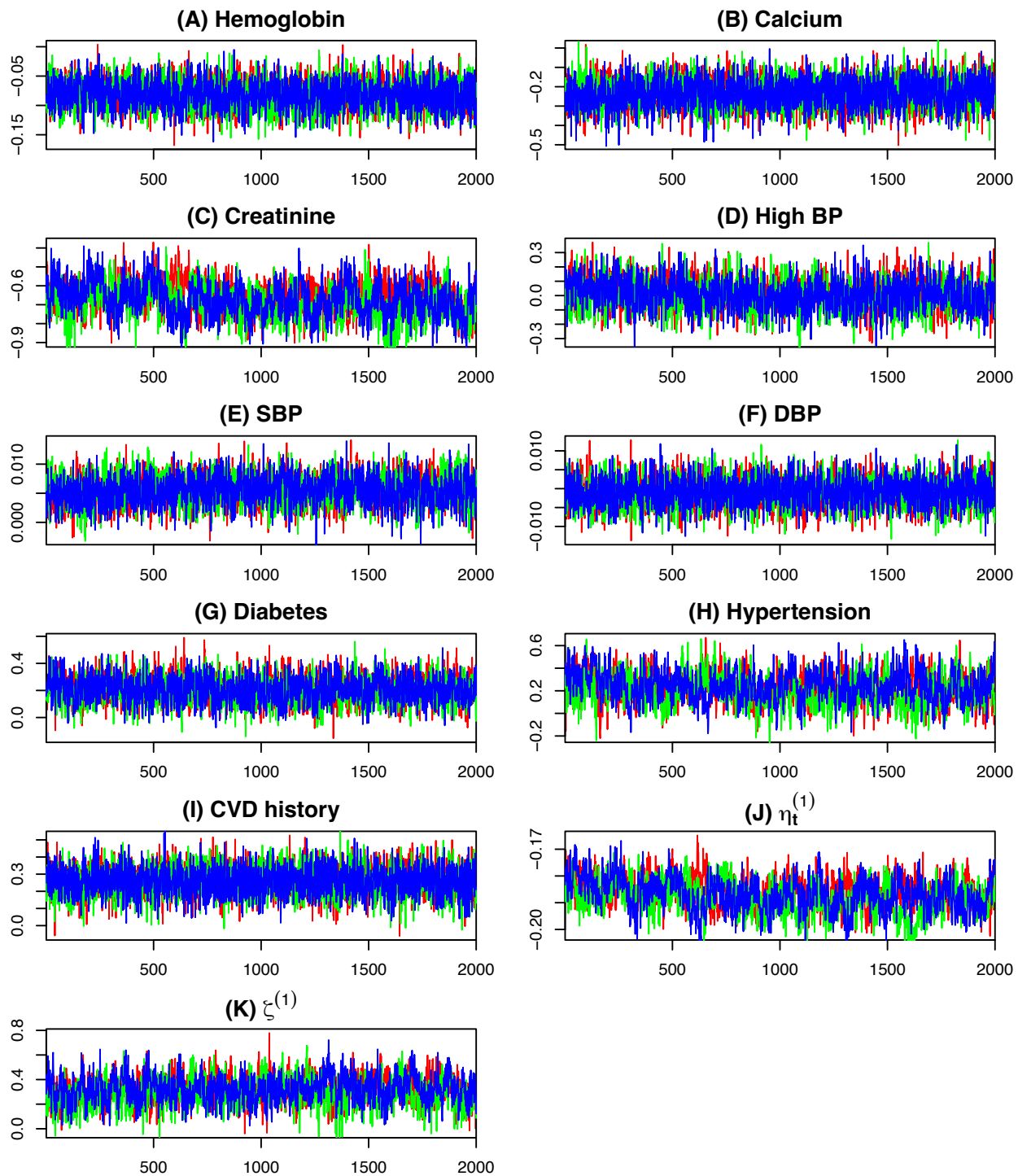


Figure S6: Trace plots for ESKD submodel parameters (part 2).

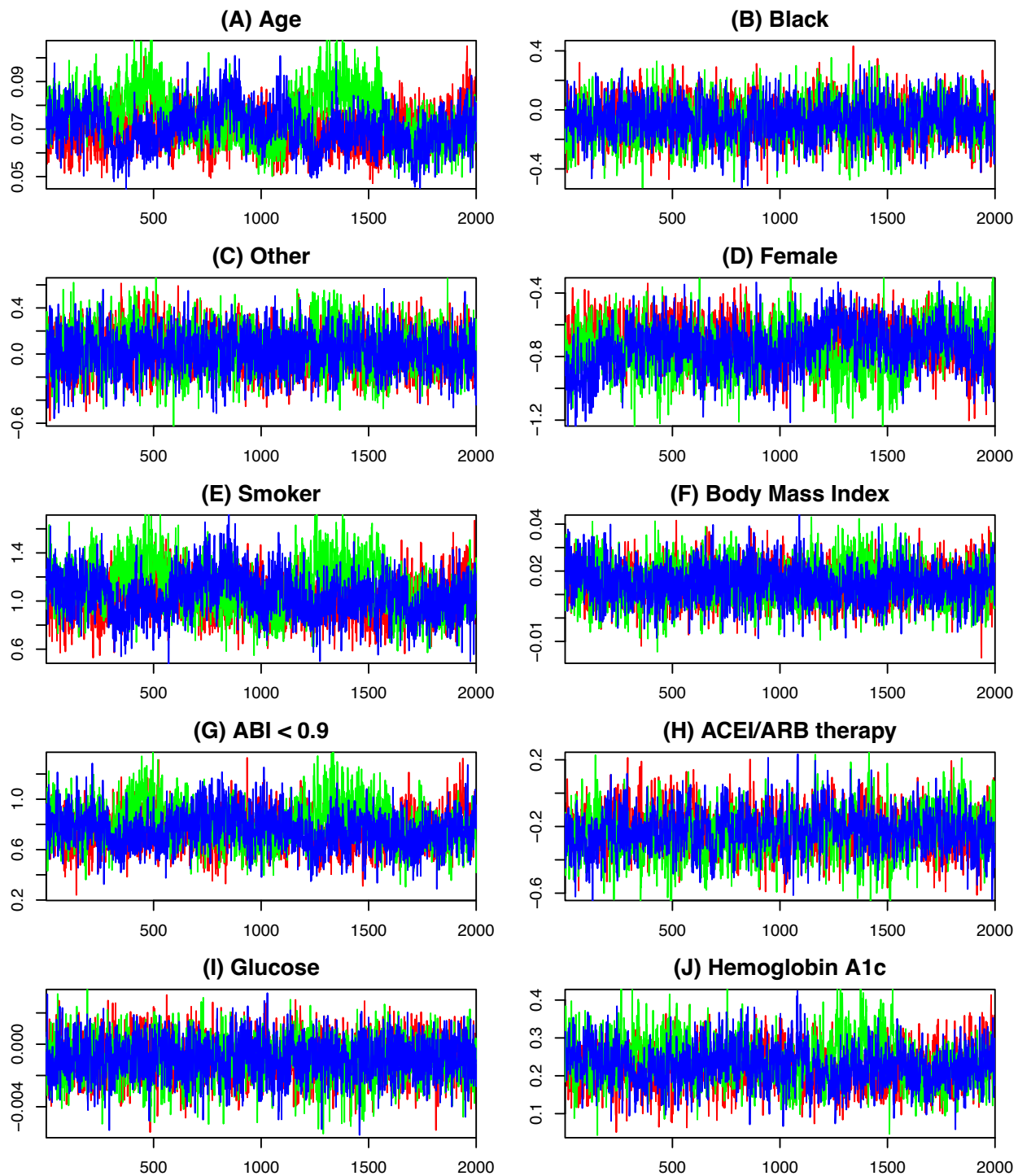


Figure S7: Trace plots for death submodel parameters (part 1).

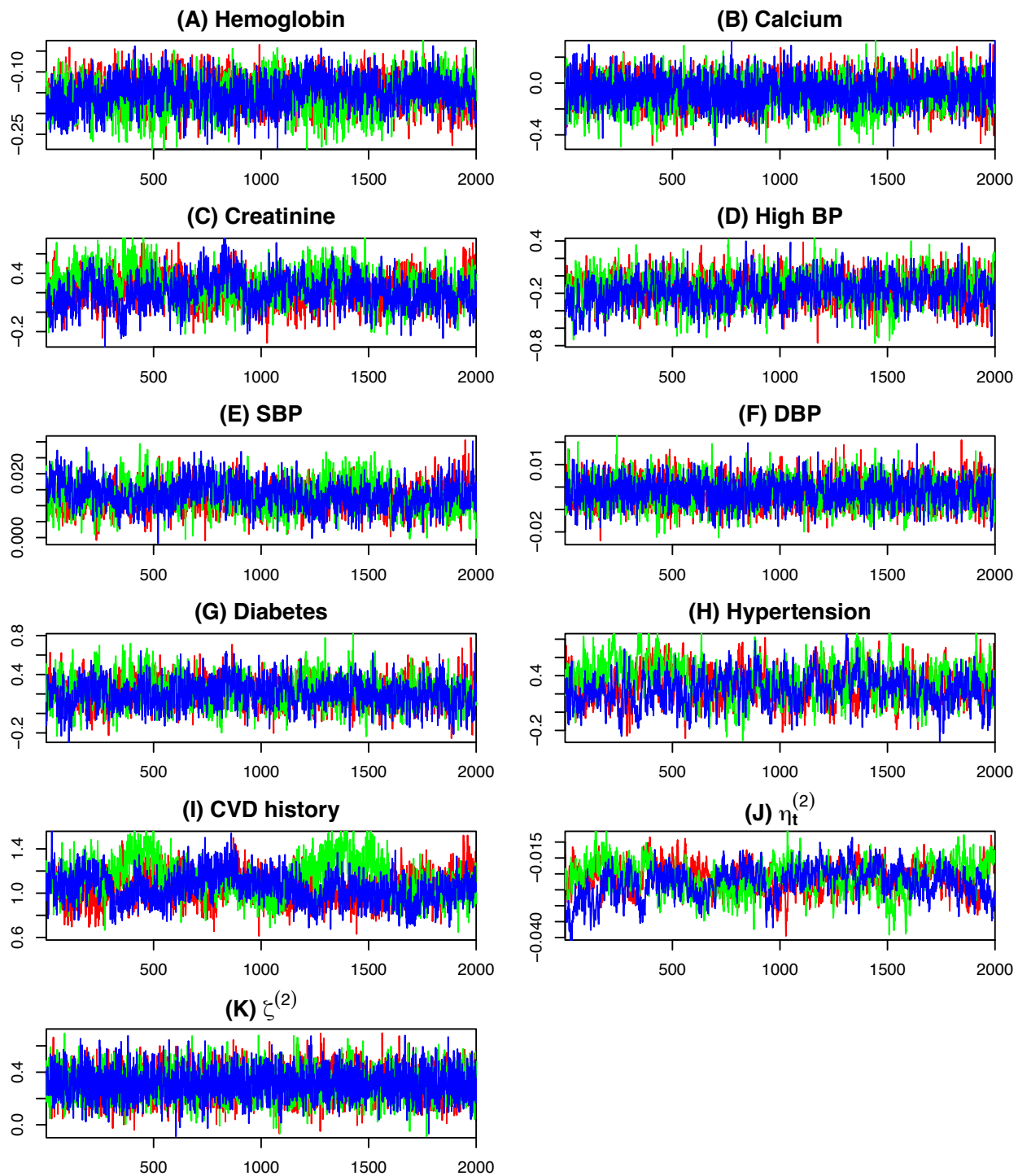


Figure S8: Trace plots for death submodel parameters (part 2).

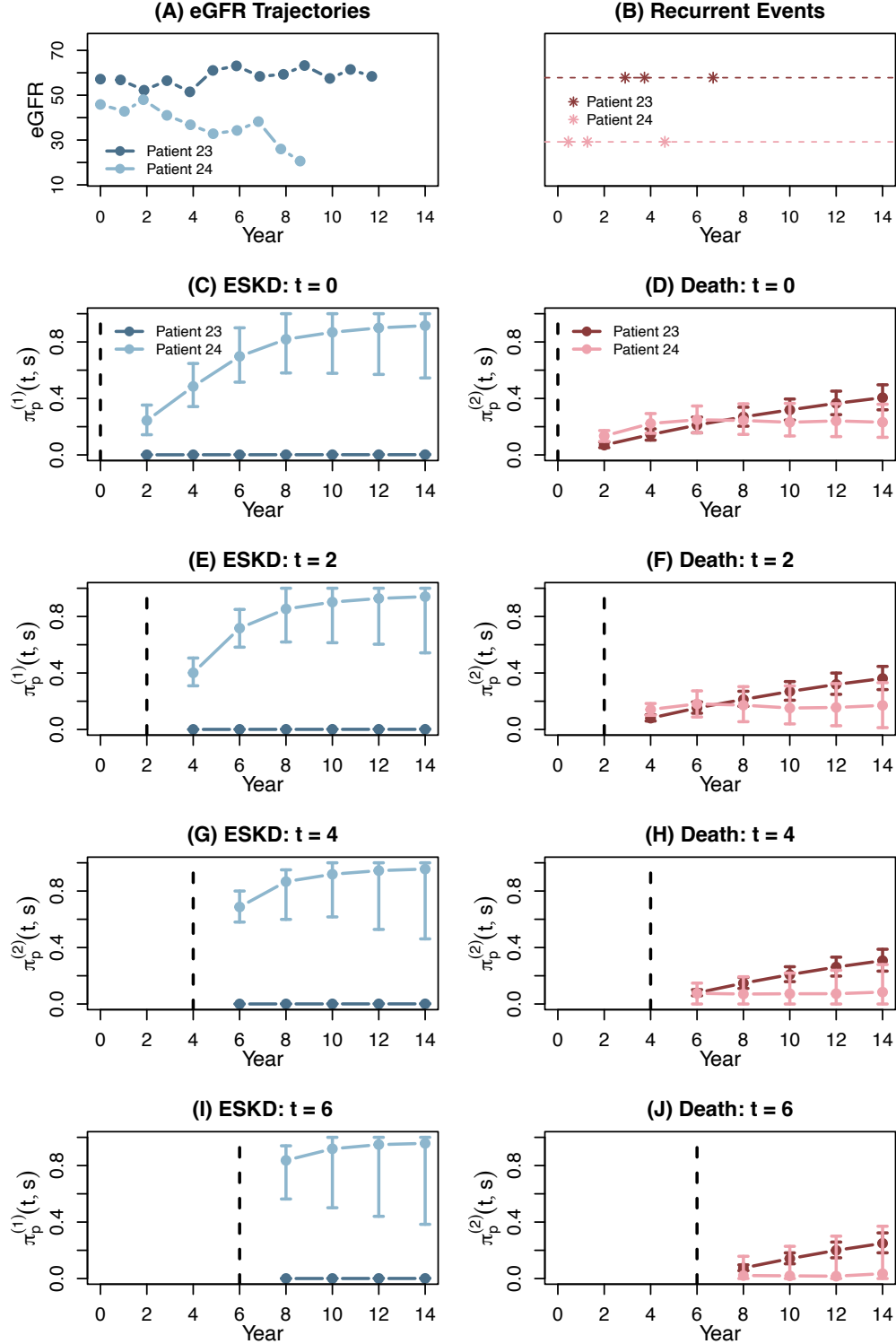


Figure S9: The eGFR trajectories (A) and recurrent CV events profiles (B), and the estimated median $\pi_p^{(w)}(t, s)$, as well as pointwise credible intervals ((C, E, G, I) for ESKD and (D, F, H, J) for death) at different prediction times, i.e., $t = 0, 2, 4, 6$ years, for Patient 23 (dark) and Patient 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.

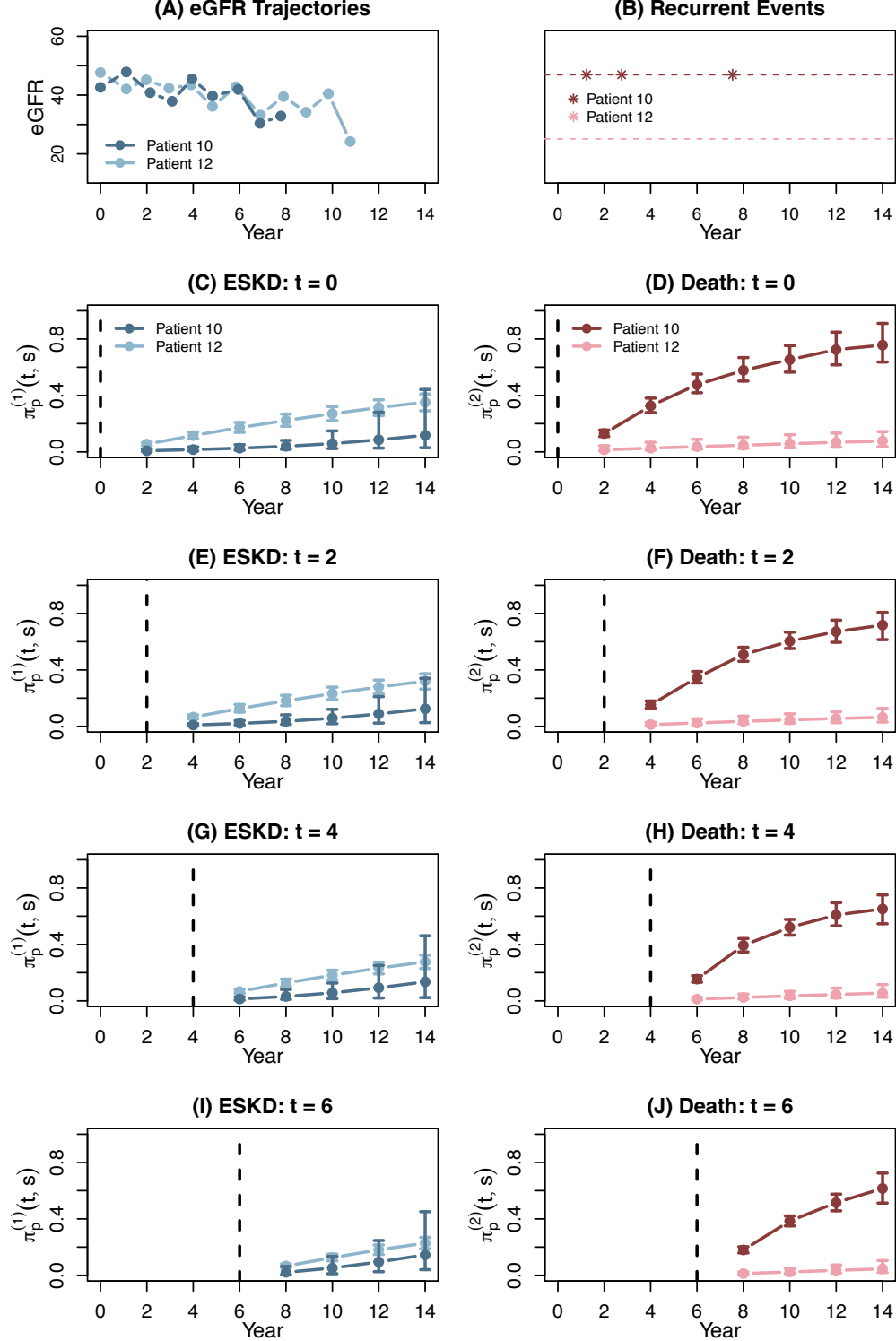


Figure S10: The eGFR trajectories (A) and recurrent CV events profiles (B), and the estimated median $\pi_p^{(w)}(t, s)$, as well as pointwise credible intervals ((C, E, G, I) for ESKD and (D, F, H, J) for death) at different prediction times, i.e., $t = 0, 2, 4, 6$ years, for Patient 10 (dark) and Patient 12 (light). The vertical dashed line denotes the time of prediction.

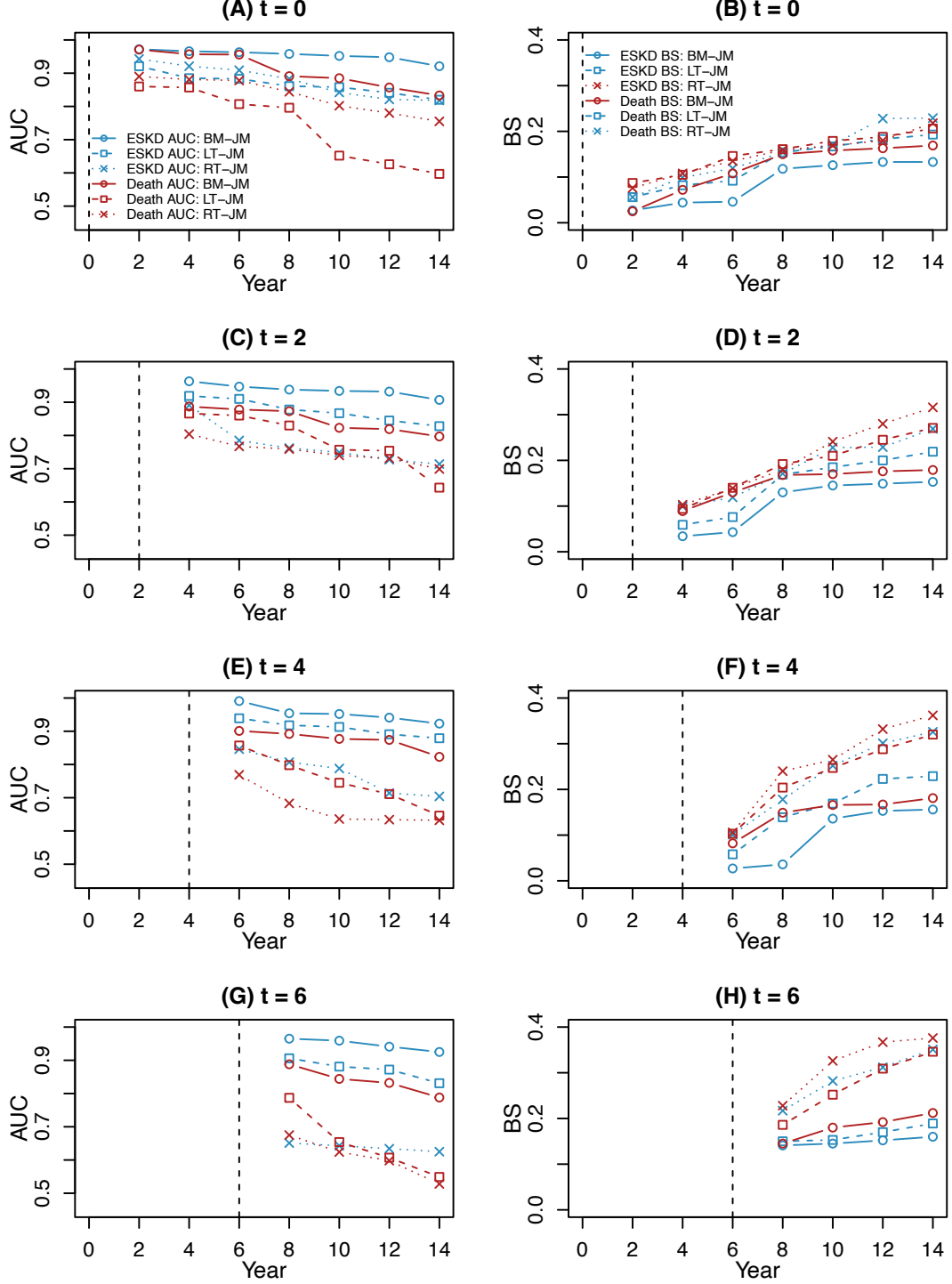


Figure S11: Dynamic AUC (A, C, E, G) and BS (B, D, F, H) values from estimated $\pi_p^{(w)}(t, s)$ for ESKD (blue) and death (red) at different prediction times, i.e., (A, B) $t = 0$, (C, D) $t = 2$ years, (E, F) $t = 4$ years, and (G, H) $t = 6$ years, until end of follow-up. The results are based on 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 S1: Monte Carlo simulation procedure for drawing $\pi_p^{(w)}(t, s)$ samples under the Bayesian formulation of the proposed BM-JM.

Algorithm 1: 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 the empirical Bayes estimates $\hat{\mathbf{u}}_p$

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

(b) Generate a proposal from a multivariate t -distribution with 4 degrees of freedom (MVT_4) $\tilde{\mathbf{u}}_p \sim \text{MVT}_4(\hat{\mathbf{u}}_p, \widehat{\text{Var}}(\hat{\mathbf{u}}_p))$, where

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

(c) Compute the Metropolis-Hastings acceptance ratio Λ

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

(d) Generate $\tilde{\Lambda} \sim \text{Unif}(0, 1)$. If $\tilde{\Lambda} < \Lambda$, accept the proposal where $\mathbf{u}_p^{(h)} \equiv \tilde{\mathbf{u}}_p$, otherwise, $\mathbf{u}_p^{(h)} \equiv \hat{\mathbf{u}}_p$.

Step 3: Calculate $\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}) = \frac{\text{CIF}_p(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})}{S_p(t, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)})}$.

Step 4: Repeat Steps 1-3 H times. The realizations $\{\pi_p^{(w)}(t, s, \mathbf{u}_p^{(h)}; \boldsymbol{\theta}^{(h)}), h = 1, \dots, H\}$ are used to estimate $\pi_p^{(w)}(t, s)$.

Table S2: 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/Ethnicity:	
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 S3: 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_{b0}^2	36.820	(34.846, 38.833)*
Var(eGFR Slope), σ_{b1}^2	5.160	(4.813, 5.527)*
Correlation(Intercept, Slope), ρ_b	0.306	(0.300, 0.321)*
Measurement Error Variance, σ_ϵ^2	37.378	(36.770, 38.013)*
σ_ν^2 (Terminal & Recurrent Link Random Effect)	0.689	(0.563, 0.835)*

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

Table S4: 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.

Variable	(A) Longitudinal eGFR	(B) Recurrent CV Events	(C1) Terminal ESKD Event	(C2) Terminal Death Event
	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)				
σ_ϵ^2	37.831 (37.199, 38.494)*	-	-	-
$\sigma_{b_0}^2$	35.506 (33.532, 37.515)*	-	-	-
$\sigma_{b_1}^2$	4.437 (4.138, 4.753)*	-	-	-
ρ_b	0.311 (0.300, 0.335)*	-	-	-
σ_ν^2	-	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. ★ 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 S5: 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)			
σ_{b0}^2	36.501 (34.433, 38.691)*	-	-
σ_{b1}^2	5.188 (4.842, 5.562)*	-	-
ρ_b	0.306 (0.300, 0.322)*	-	-
σ_ϵ^2	37.406 (36.774, 38.062)*	-	-
$\eta_t^{(1)}$	-	-0.18 (-0.200, -0.176)*	-
$\eta_t^{(2)}$	-	-	-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 S6: 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.

Variable	(A) Recurrent CV Events	(B1) Terminal ESKD Event	(B2) Terminal Death Event
	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)			
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 S7: 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.

	True	BM-JM			LT-JM			RT-JM		
		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										
$\sigma_{b_0}^2$	1.25	-0.033	7.0×10^{-4}	94.5	-0.045	1.3×10^{-3}	93.5	-	-	-
$\sigma_{b_1}^2$	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_ε^2	1.32	0.004	9.2×10^{-6}	95.0	0.005	1.4×10^{-5}	93.0	-	-	-
σ_ν^2	1.44	-0.167	0.013	90.0	-	-	-	1.231	0.731	0

Table S8: 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.

	True	L-M			R-M			T-M		
		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										
$\sigma_{b_0}^2$	1.25	-0.037	8.9×10^{-4}	86.0	-	-	-	-	-	-
$\sigma_{b_1}^2$	0.80	-0.142	0.032	59.5	-	-	-	-	-	-
ρ_b	0.50	-0.102	0.042	47.5	-	-	-	-	-	-
σ_ε^2	1.32	0.007	3.0×10^{-5}	92.0	-	-	-	-	-	-
σ_ν^2	1.44	-	-	-	1.087	0.570	0	-0.250	0.030	66.5

Table S9: 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 $w = 1, 2$.

t	Δt	AUC $_{\pi^{(1)}}$			AUC $_{\pi^{(2)}}$			BS $_{\pi^{(1)}}$			BS $_{\pi^{(2)}}$		
		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