

# **Supplementary Material for “Combining Dynamic Predictions from Joint Models for Longitudinal and Time-to-Event Data using Bayesian Model Averaging”**

## **1 Aortic Valve Dataset**

This section presents figures and tables for the analysis of the Aortic Valve dataset presented in Section 6 of the main manuscript. More specifically:

- Table 1 presents estimates and 95% credibility intervals for the parameters of the longitudinal submodel under the five association structures.
- Table 2 presents estimates and 95% credibility intervals for the parameters of the survival submodel under the five association structures.
- Figures 1 depicts the predicted re-operation-free survival probabilities for Patients 22 under the five joint models along with the BMA predictions.
- Figures 2–4 depict predicted square root aortic gradient levels for Patients 20, 22 and 81, respectively, under the five joint models along with the BMA predictions.

	Value ( $M_1$ )	95% CI	Value+Slope ( $M_2$ )	95% CI	Area ( $M_3$ )	95% CI	Weighted Area ( $M_4$ )	95% CI	Shared RE ( $M_5$ )	Est.	95% CI
$S_{\text{I}}$	3.237	(2.940; 3.553)	3.216	(2.899; 3.537)	3.235	(2.929; 3.554)	3.258	(2.944; 3.563)	3.229	(2.912; 3.535)	
$R_{\text{R}}$	2.831	(2.695; 2.969)	2.831	(2.688; 2.975)	2.835	(2.693; 2.970)	2.839	(2.696; 2.984)	2.833	(2.693; 2.979)	
$S_{\text{I}}:\text{B-spln1}$	0.969	(0.566; 1.370)	0.996	(0.592; 1.426)	0.940	(0.546; 1.331)	0.923	(0.530; 1.321)	1.073	(0.641; 1.526)	
$R_{\text{R}}:\text{B-spln1}$	0.993	(0.652; 1.337)	1.013	(0.667; 1.353)	1.003	(0.658; 1.344)	1.003	(0.673; 1.345)	1.127	(0.786; 1.491)	
$S_{\text{I}}:\text{B-spln2}$	2.479	(1.647; 3.276)	2.562	(1.798; 3.368)	2.437	(1.640; 3.221)	2.370	(1.581; 3.137)	2.897	(1.994; 3.835)	
$R_{\text{R}}:\text{B-spln2}$	1.905	(1.279; 2.578)	1.933	(1.209; 2.567)	1.760	(1.126; 2.393)	1.712	(1.059; 2.391)	2.123	(1.397; 2.858)	
$S_{\text{I}}:\text{B-spln3}$	2.724	(1.982; 3.510)	2.791	(2.072; 3.600)	2.614	(1.937; 3.372)	2.582	(1.911; 3.284)	3.341	(2.424; 4.346)	
$R_{\text{R}}:\text{B-spln3}$	1.848	(0.840; 2.877)	1.883	(0.815; 2.921)	1.607	(0.625; 2.608)	1.527	(0.564; 2.555)	2.149	(1.061; 3.293)	
$\sigma$	0.352	(0.326; 0.380)	0.354	(0.328; 0.380)	0.355	(0.330; 0.383)	0.356	(0.329; 0.384)	0.359	(0.333; 0.387)	
$D_{11}$	0.428	(0.311; 0.576)	0.417	(0.308; 0.558)	0.433	(0.312; 0.585)	0.432	(0.315; 0.577)	0.423	(0.306; 0.569)	
$D_{21}$	-0.016	(-0.198; 0.167)	-0.011	(-0.183; 0.166)	0.009	(-0.173; 0.186)	-0.015	(-0.202; 0.163)	-0.010	(-0.187; 0.169)	
$D_{31}$	-0.151	(-0.513; 0.151)	-0.145	(-0.496; 0.158)	-0.128	(-0.493; 0.163)	-0.186	(-0.588; 0.128)	-0.175	(-0.560; 0.163)	
$D_{41}$	0.071	(-0.345; 0.475)	0.035	(-0.356; 0.432)	0.106	(-0.278; 0.471)	0.019	(-0.381; 0.395)	0.018	(-0.401; 0.452)	
$D_{22}$	1.162	(0.648; 1.832)	1.203	(0.699; 1.869)	1.063	(0.589; 1.669)	1.043	(0.589; 1.651)	1.194	(0.661; 1.902)	
$D_{32}$	1.718	(1.052; 2.545)	1.855	(1.198; 2.707)	1.517	(0.930; 2.264)	1.558	(0.957; 2.328)	1.966	(1.220; 2.893)	
$D_{42}$	1.600	(0.811; 2.596)	1.742	(0.950; 2.762)	1.374	(0.690; 2.190)	1.396	(0.726; 2.180)	1.817	(1.007; 2.836)	
$D_{33}$	3.747	(2.125; 5.938)	4.027	(2.402; 6.371)	3.406	(1.910; 5.408)	3.660	(2.137; 5.667)	4.617	(2.671; 7.137)	
$D_{43}$	3.180	(1.339; 5.708)	3.540	(1.690; 6.433)	2.719	(1.106; 4.941)	2.922	(1.351; 4.983)	4.015	(1.873; 6.890)	
$D_{44}$	4.006	(1.901; 7.334)	4.330	(2.053; 8.122)	3.329	(1.483; 6.006)	3.353	(1.684; 5.828)	4.586	(2.129; 8.111)	

Table 1: Estimated coefficients and 95% credibility intervals for the parameters of the longitudinal submodels, under the five joint models fitted to the Aortic Valve dataset.  $D_{ij}$  denotes the  $ij$ -th element of the covariance matrix of the random effects  $\mathbf{D}$ .

	Value ( $M_1$ )	95% CI	Value+Slope ( $M_2$ )	95% CI	Est.	Area ( $M_3$ )	95% CI	Weighted Area ( $M_4$ )	95% CI	Shared RE ( $M_5$ )	95% CI
RR	0.322	(-0.071; 0.719)	0.338	(-0.059; 0.746)	0.212	(-0.178; 0.617)	0.088	(-0.317; 0.493)	0.150	(-0.395; 0.703)	
Age	0.004	(-0.009; 0.016)	0.007	(-0.006; 0.020)	-0.003	(-0.015; 0.009)	-0.003	(-0.015; 0.009)	0.013	(-0.003; 0.029)	
Female	-0.191	(-0.586; 0.188)	-0.160	(-0.570; 0.240)	-0.144	(-0.547; 0.227)	-0.092	(-0.501; 0.296)	-0.138	(-0.676; 0.346)	
$\alpha_1$	0.332	(0.191; 0.467)	0.217	(0.051; 0.384)	0.015	(-0.003; 0.033)	-0.018	(-0.049; 0.013)	-0.351	(-1.234; 0.229)	
$\alpha_2$									-0.131	(-1.807; 1.306)	
$\alpha_3$									-0.010	(-0.573; 0.686)	
$\alpha_4$									0.479	(-0.059; 1.297)	
$\gamma_{h_0,1}$	-5.898	(-7.622; -4.346)	-6.270	(-8.192; -4.575)	-4.584	(-6.283; -3.169)	-4.465	(-6.127; -3.028)	-6.725	(-9.640; -4.612)	
$\gamma_{h_0,2}$	-5.883	(-7.817; -3.972)	-6.081	(-8.079; -4.144)	-4.421	(-6.249; -2.651)	-4.131	(-5.978; -2.353)	-6.134	(-8.317; -4.099)	
$\gamma_{h_0,3}$	-4.781	(-6.445; -3.127)	-4.700	(-6.387; -3.040)	-3.386	(-4.908; -1.887)	-2.795	(-4.407; -1.206)	-4.458	(-6.250; -2.758)	
$\gamma_{h_0,4}$	-4.948	(-6.267; -3.626)	-4.847	(-6.214; -3.505)	-3.592	(-4.790; -2.426)	-2.632	(-3.926; -1.310)	-4.179	(-5.427; -2.971)	
$\gamma_{h_0,5}$	-4.180	(-5.435; -2.882)	-3.994	(-5.286; -2.719)	-2.932	(-4.231; -1.722)	-1.750	(-3.229; -0.339)	-3.191	(-4.347; -2.029)	
$\gamma_{h_0,6}$	-4.650	(-6.067; -3.166)	-4.447	(-5.878; -3.025)	-3.461	(-4.937; -2.013)	-1.978	(-3.618; -0.281)	-3.368	(-4.671; -2.113)	
$\gamma_{h_0,7}$	-2.988	(-4.936; -1.079)	-2.722	(-4.709; -0.795)	-1.967	(-4.182; 0.083)	-0.092	(-2.513; 2.230)	-1.399	(-3.309; 0.519)	
$\gamma_{h_0,8}$	-5.539	(-8.379; -2.749)	-5.289	(-8.223; -2.495)	-4.628	(-7.771; -1.593)	-2.316	(-5.697; 1.005)	-3.605	(-6.556; -0.820)	
$\gamma_{h_0,9}$	-3.882	(-7.759; -0.784)	-3.640	(-7.510; -0.542)	-3.495	(-7.644; -0.229)	-1.672	(-6.280; 2.128)	-2.267	(-6.430; 1.245)	

Table 2: Estimated coefficients and 95% credibility intervals for the parameters of the survival submodels, under the five joint models fitted to the Aortic Valve dataset.

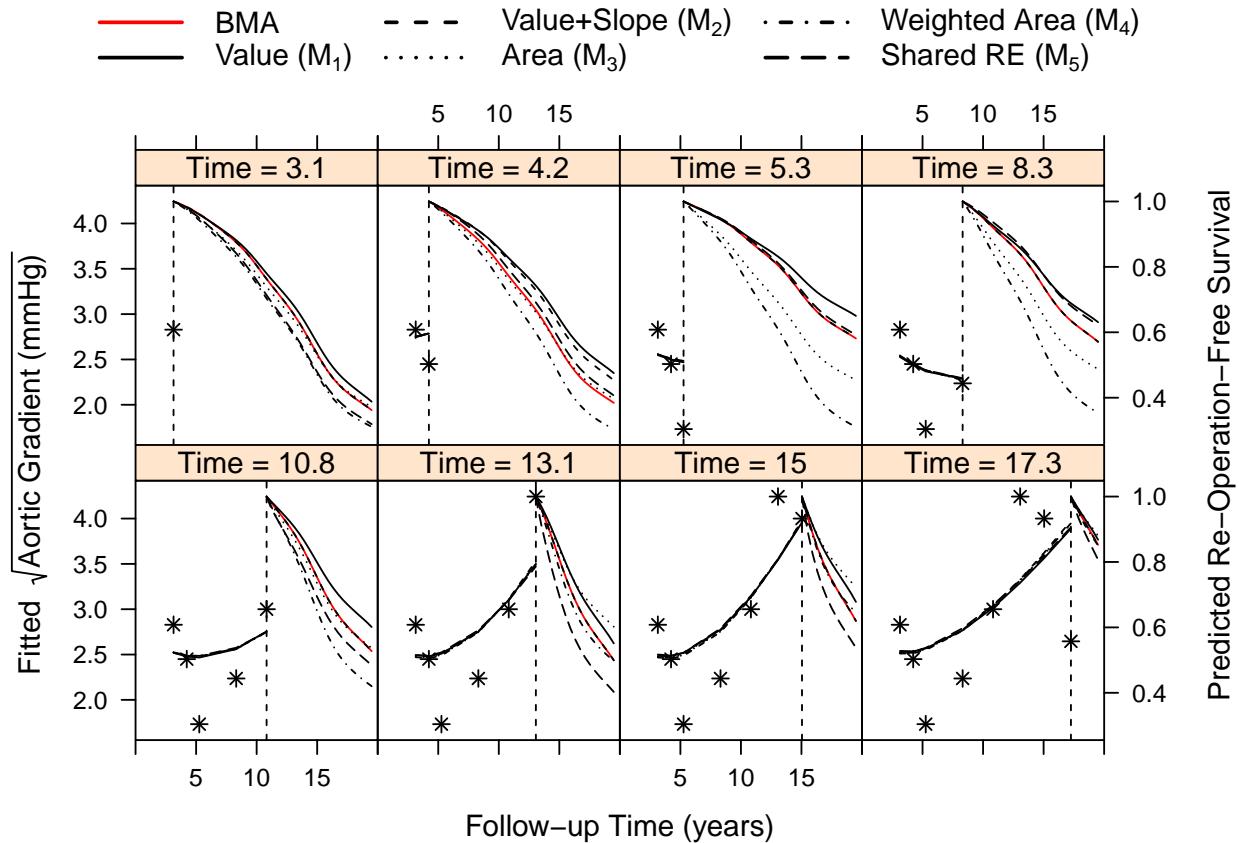


Figure 1: Dynamic predictions of re-operation-free survival for Patient 22 under the five joint models along with the BMA predictions. Each panel shows the corresponding conditional survival probabilities calculated after each of his longitudinal measurements have been recorded.

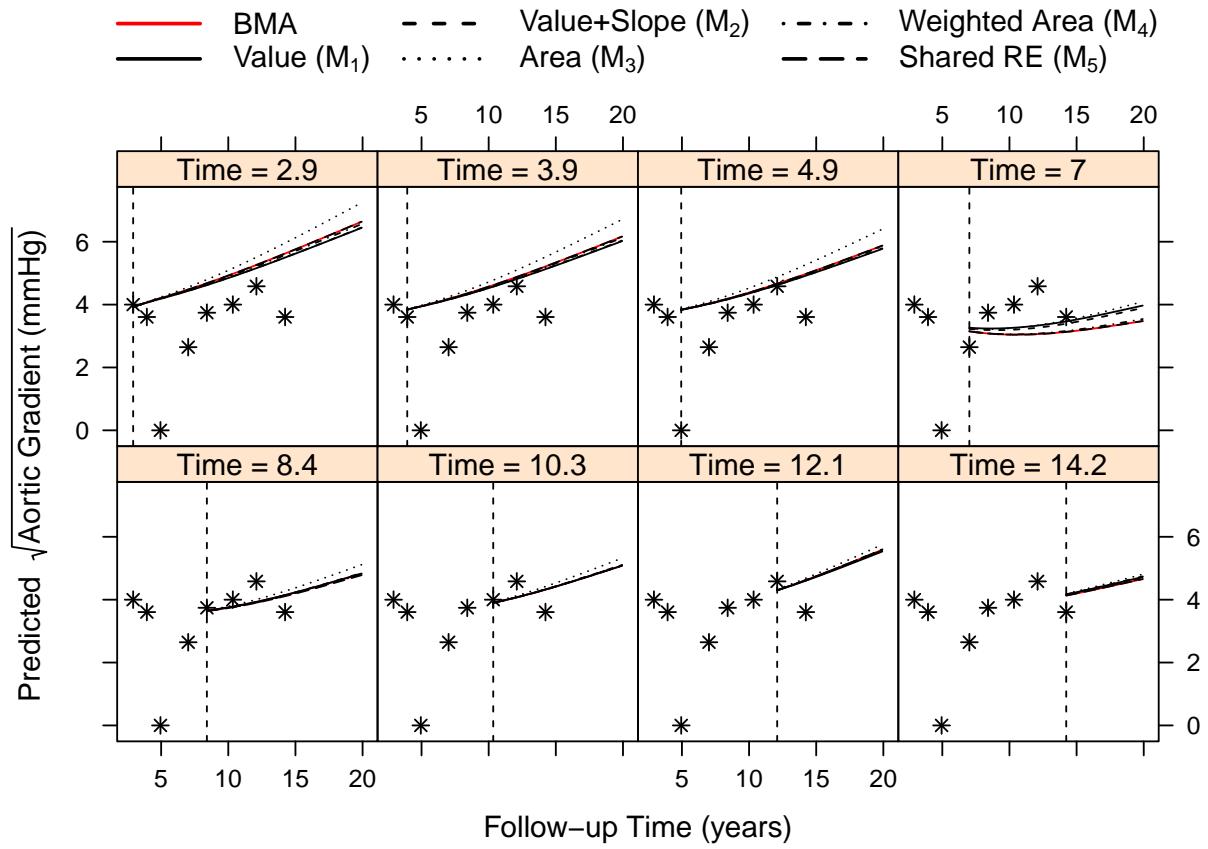


Figure 2: Dynamic predictions of square root aortic gradient levels for Patient 20 under the five joint models along with the BMA predictions. Each panel shows the corresponding predictions calculated after each of her longitudinal measurements have been recorded, denoted by the vertical dashed line (i.e., measurements before the vertical line are used in the calculation of predictions). Asterisks denote the observed square root aortic gradient levels.

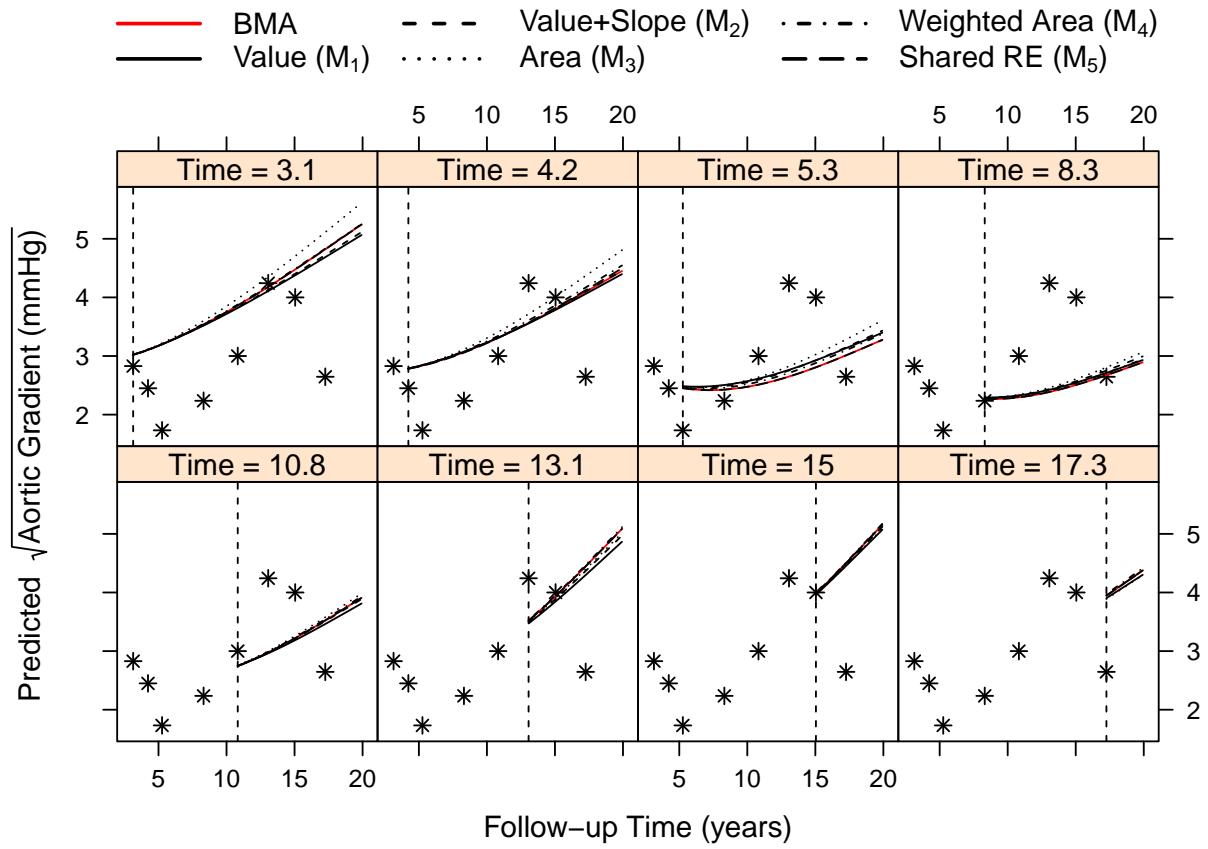


Figure 3: Dynamic predictions of square root aortic gradient levels for Patient 22 under the five joint models along with the BMA predictions. Each panel shows the corresponding predictions calculated after each of his longitudinal measurements have been recorded, denoted by the vertical dashed line (i.e., measurements before the vertical line are used in the calculation of predictions). Asterisks denote the observed square root aortic gradient levels.

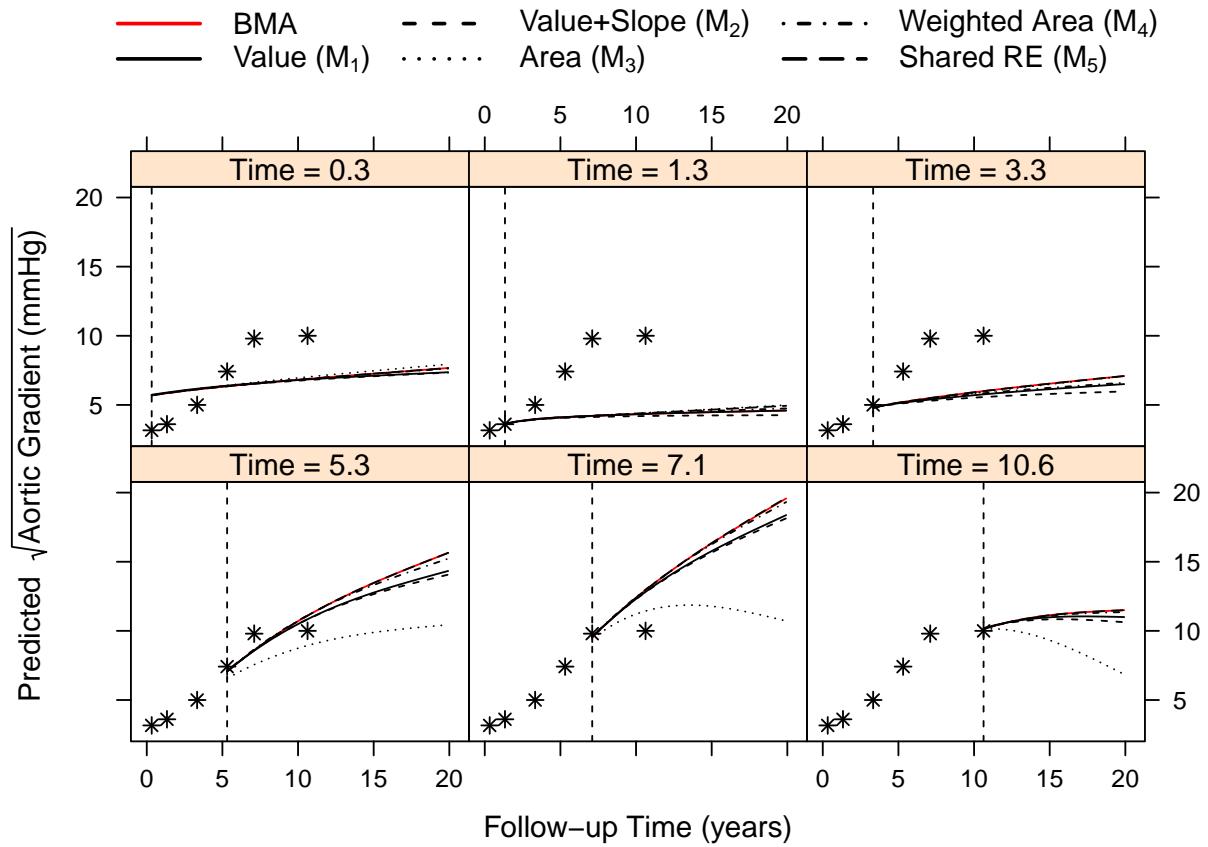


Figure 4: Dynamic predictions of square root aortic gradient levels for Patient 81 under the five joint models along with the BMA predictions. Each panel shows the corresponding predictions calculated after each of his longitudinal measurements have been recorded, denoted by the vertical dashed line (i.e., measurements before the vertical line are used in the calculation of predictions). Asterisks denote the observed square root aortic gradient levels.

## 2 Laplace Approximations

In this section we describe the details behind the Laplace approximations performed to evaluate the marginal densities  $p(\mathcal{D}_n \mid M_k)$  and  $p(\mathcal{D}_j(t) \mid M_k)$ , presented in Section 6 of the main paper. Because both of them require very similar calculations, we only show here the Laplace approximation for  $p(\mathcal{D}_n \mid M_k)$  with  $p(\mathcal{D}_j(t) \mid M_k)$  approximated analogously. In addition, to facilitate the use of the Laplace approximation, in the following  $\boldsymbol{\theta}$  denotes the ‘unconstrained’ parameters, namely  $\log(\sigma)$  instead of  $\sigma$ , and with the Cholesky factor of  $\mathbf{D}$  instead of  $\mathbf{D}$ . Based on the notation and definitions of Section 2 of the main paper,  $p(\mathcal{D}_n \mid M_k)$  is written as

$$\begin{aligned} p(\mathcal{D}_n \mid M_k) &= \prod_{i=1}^n p(\mathbf{y}_i, T_i, \delta_i \mid M_k) \\ &= \prod_{i=1}^n \int \left[ \int \left\{ \prod_{l=1}^{n_i} p(y_{il} \mid \mathbf{b}_i, \boldsymbol{\theta}_k) \right\} p(T_i, \delta_i \mid \mathbf{b}_i, \boldsymbol{\theta}_k) p(\mathbf{b}_i \mid \boldsymbol{\theta}_k) d\mathbf{b}_i \right] p(\boldsymbol{\theta}_k \mid M_k) d\boldsymbol{\theta}_k. \end{aligned} \quad (1)$$

We let

$$\begin{aligned} \{\hat{\boldsymbol{\theta}}^\top, \hat{\mathbf{b}}_i^\top\}^\top &= \arg \max_{\boldsymbol{\theta}, \mathbf{b}_i} \left\{ \prod_i p(\boldsymbol{\theta}, \mathbf{b} \mid \mathbf{y}_i, T_i, \delta_i, M_k) \right\} \\ &= \arg \max_{\boldsymbol{\theta}, \mathbf{b}_i} \left\{ \sum_i \log p(\mathbf{y}_i \mid \mathbf{b}, \boldsymbol{\theta}) + \log p(T_i, \delta_i \mid \mathbf{b}, \boldsymbol{\theta}) + \log p(\mathbf{b} \mid \boldsymbol{\theta}) + \log p(\boldsymbol{\theta} \mid M_k) \right\}, \end{aligned}$$

denote the mode of the full posterior distribution of the parameters and random effects. This can be obtained by either a direct optimization of the joint log-likelihood function or using the MCMC sample at hand (e.g., by applying a density estimation approach and finding the mode for each parameter and random effect). Let also

$$\Sigma_{\mathbf{b}_i} = - \frac{\partial^2 \log p(\mathbf{y}_i \mid \mathbf{b}, \hat{\boldsymbol{\theta}}_k) + \log p(T_i, \delta_i \mid \mathbf{b}, \hat{\boldsymbol{\theta}}_k) + \log p(\mathbf{b} \mid \hat{\boldsymbol{\theta}}_k) + \log p(\hat{\boldsymbol{\theta}}_k \mid M_k)}{\partial \mathbf{b}^\top \partial \mathbf{b}} \Big|_{\mathbf{b}=\hat{\mathbf{b}}_i},$$

denote the Hessian matrix for the random effects, and analogously,

$$\Sigma_\theta = -\frac{\partial^2 \sum_i \left\{ \log p(\mathbf{y}_i | \hat{\mathbf{b}}_i, \boldsymbol{\theta}) + \log p(T_i, \delta_i | \hat{\mathbf{b}}_i, \boldsymbol{\theta}) + \log p(\hat{\mathbf{b}}_i | \boldsymbol{\theta}) + \log p(\boldsymbol{\theta} | M_k) \right\}}{\partial \boldsymbol{\theta}^\top \partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},$$

denote the Hessian matrix for the parameters. Then, we approximate the inner integral in (1) by

$$\begin{aligned} & p(\mathbf{y}_i, T_i, \delta_i | \hat{\boldsymbol{\theta}}) \\ & \approx \exp \left[ \frac{q \log(2\pi) - \log \{ \det(\Sigma_{b_i}) \}}{2} + \log p(\mathbf{y}_i | \hat{\mathbf{b}}_i, \hat{\boldsymbol{\theta}}) + \log p(T_i, \delta_i | \hat{\mathbf{b}}_i, \hat{\boldsymbol{\theta}}) + \log p(\hat{\mathbf{b}}_i | \hat{\boldsymbol{\theta}}) \right], \end{aligned}$$

where  $q$  denotes the number of random effects for each subject  $i$ . Similarly, the outer integral in (1) is approximated as

$$p(\mathcal{D}_n | M_k) \approx \exp \left[ \frac{\kappa \log(2\pi) - \log \{ \det(\Sigma_\theta) \}}{2} + \sum_i \log p(\mathbf{y}_i, T_i, \delta_i | \hat{\boldsymbol{\theta}}) \right],$$

where  $\kappa$  denotes the number of parameters.

## 3 Simulation

### 3.1 Simulation Settings

For all simulation scenarios the parameter values that were used for the longitudinal submodel were

Fixed effects:  $\beta_1 = 3.35$ ,  $\beta_2 = 2.84$ ,  $\beta_3 = 1.51$ ,  $\beta_4 = 1.25$ ,  $\beta_5 = 2.47$ ,  $\beta_6 = 2.30$ ,  $\beta_7 = 2.54$ , and  $\beta_8 = 1.85$ ;

Random effects covariance matrix:

$$\mathbf{D} = \begin{bmatrix} 0.43 \\ -0.02 & 1.16 \\ -0.15 & 1.72 & 3.75 \\ 0.07 & 1.60 & 3.18 & 4.01 \end{bmatrix}$$

Measurement error standard deviation:  $\sigma = 0.60$ .

For the survival submodels the parameters that were used to simulate from each scenario are given in Table 3.

	Scenario			
	I	II	III	IV
$\gamma_0$	-6.73	-6.73	-16.73	-4.93
$\gamma_1$	0.41	0.41	0.41	0.41
$\alpha_1$	0.39	0.66	0.30	-0.55
$\alpha_2$		-2.85		-1.00
$\alpha_3$				0.35
$\alpha_4$				0.90
$\sigma_t$	1.65	1.65	1.65	1.60

Table 3: Parameter values for the survival submodels under the four simulation scenarios.

### 3.2 Extra Results

Figure 5 shows boxplots of the distribution of  $\log p(\mathcal{D}_n \mid M_k)$  for the five joint models fitted in each scenario:

$$\begin{aligned}
M_1 \quad h_i(t) &= h_0(t) \exp\{\gamma_0 + \gamma_1 \text{Trt} \mathbf{1}_i + \alpha_1 m_i(t)\}, \\
M_2 \quad h_i(t) &= h_0(t) \exp\{\gamma_0 + \gamma_1 \text{Trt} \mathbf{1}_i + \alpha_1 m_i(t) + \alpha_2 m'_i(t)\}, \\
M_3 \quad h_i(t) &= h_0(t) \exp\{\gamma_0 + \gamma_1 \text{Trt} \mathbf{1}_i + \alpha_1 \int_0^t m_i(s) ds\}, \\
M_4 \quad h_i(t) &= h_0(t) \exp\{\gamma_0 + \gamma_1 \text{Trt} \mathbf{1}_i + \alpha_1 \int_0^t \frac{\phi(t-s)}{\Phi(t) - 0.5} m_i(s) ds\}, \\
M_5 \quad h_i(t) &= h_0(t) \exp(\gamma_0 + \gamma_1 \text{Trt} \mathbf{1}_i + \alpha_1 b_{i0} + \alpha_2 b_{i1} + \alpha_3 b_{i2} + \alpha_4 b_{i3}),
\end{aligned}$$

where  $\phi(\cdot)$  and  $\Phi(\cdot)$  denote the probability density and cumulative distribution functions of the normal distribution with mean zero and standard deviation 1/3.

In Scenario I model  $M_1$  is the true model, in Scenario II model  $M_2$  is the true model, in Scenario III model  $M_3$  is the true model, and in Scenario IV model  $M_5$  is the true model.

Figure 6 shows the distribution of the ranks of the root mean squared error of the true model, of BMA including the true model and BMA without the true model in the four simulation scenarios.

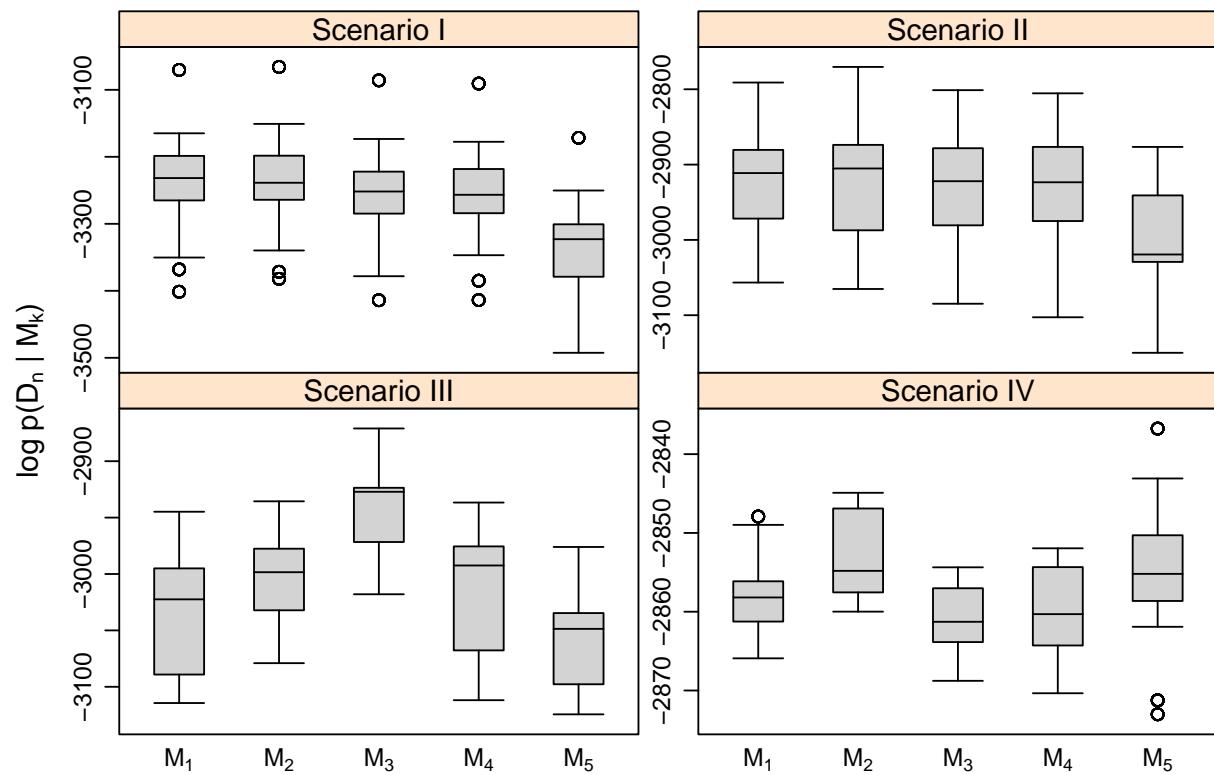


Figure 5: Simulation results under the four scenarios based on 200 datasets. Each boxplot shows the distribution of  $\log p(\mathcal{D}_n \mid M_k)$  for each of the five fitted joint models, under each scenario.

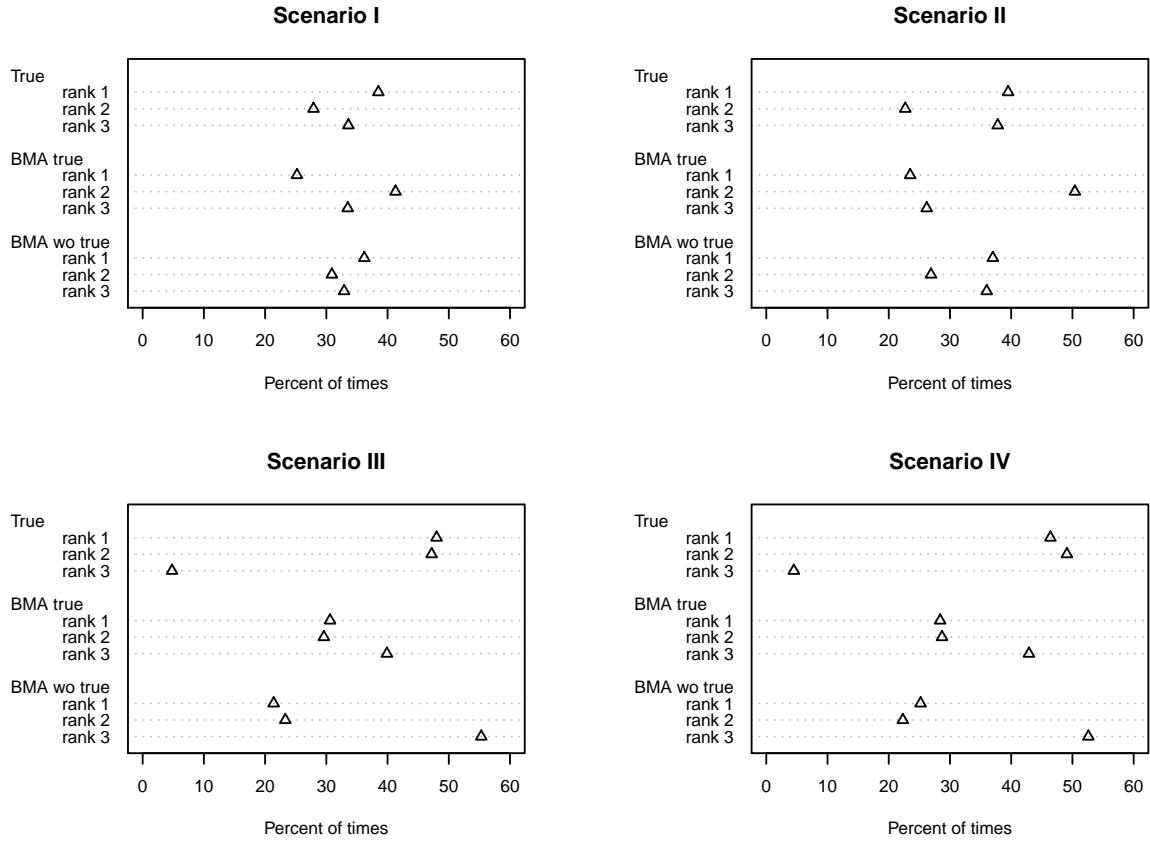


Figure 6: Simulation results under the four scenarios based on 200 datasets. Each dotchart shows the distribution of the ranks of the root mean squared error of the true model, of BMA including the true model and BMA without the true model versus the gold standard based on the true values for the parameters and the random effects.