Supplementary Materials for "Bayesian kernel machine regression–causal mediation analysis" by Katrina L. Devick, Jennifer F. Bobb, Maitreyi M. Mazumdar, Birgit Claus Henn, David C. Bellinger, David C. Christiani, Robert O. Wright, Paige L. Williams, Brent A. Coull, and Linda Valeri

Section A: Details

Section A.1: Prior specification

Here we specify the prior distributions for the BKMR models described in Section 2 and fit in Sections 3 and 4 of the main text. We assumed a flat prior on the coefficients for the confounding variables, $\boldsymbol{\beta} \sim 1$, and assumed $\sigma^{-2} \sim Gamma(a_\sigma, b_\sigma)$, where we set both the shape parameter a_σ and the scale parameter b_σ to 0.001. For convenience, we parameterized BKMR model (1) with $\lambda = \tau\sigma^{-2}$, where we assumed a Gamma prior distribution for λ with mean $\mu_\lambda = 10$ and variance $\sigma^2_\lambda = 100$.

When BKMR models were fit with component-wise variable selection, we assumed the distribution of the slab $f_1(r_\ell)$ in (4) of the main text was a inverse uniform, such that $\rho_\ell = \frac{1}{r_\ell} \sim Unif(a_r, b_r)$, with $a_r = 0$ and $b_r = 100$. Additionally, we assumed the prior probability π that a mixture component z_ℓ was included in the model followed a beta distribution, such that $\pi \sim Beta(a_\pi, b_\pi)$. We used $a_\pi = 2$ and $b_\pi = 6$, such that approximately 25% of the mixture components would be included in the model.

The models used for our Bayesian Kernel Machine Regression–Causal Mediation Analysis (BKMR-CMA) algorithm as well as the priors and notation are summarized in the following figure.

Summary of Bayesian Kernel Machine Regression–Causal Mediation Analysis

General BKMR model specification and prior distributions

$$\text{Likelihood}\left\{\begin{array}{ll} \mathbf{Y} \mid \mathbf{h}, \boldsymbol{\beta}, \sigma^2, \mathbf{C} & \sim \mathrm{N}(\mathbf{h} + \mathbf{C}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n) \\ \mathbf{h} \mid \tau, \mathbf{r}, \mathbf{Z} & \sim \mathrm{N}(\mathbf{0}, \tau \mathbf{K}_{\mathbf{Z},\mathbf{r}}) \end{array}\right.$$

$$\text{Component-wise variable selection}\left\{\begin{array}{ll} r_\ell \mid \delta_\ell & \sim \delta_\ell \mathrm{Unif}^{-1}(a_r, b_r) + (1-\delta_\ell) P_0, \\ \delta_\ell \mid \pi & \sim \mathrm{Bernoulli}(\pi), \end{array}\right.$$

$$\text{Priors}\left\{\begin{array}{rl} \boldsymbol{\beta} & \sim 1 \\ \sigma^{-2} & \sim \mathrm{Gamma}(a_\sigma, b_\sigma) \\ \lambda \equiv \tau\sigma^{-2} & \sim \mathrm{Gamma}(a_\lambda, b_\lambda) \\ \pi & \sim \mathrm{Beta}(a_\pi, b_\pi) \end{array}\right.$$

BKMR-CMA model specification

$$\text{Mediator model}\left\{\begin{array}{l} M_i \;=\; h_M(\mathbf{Z}_{Mi}) + \mathbf{C}_i^T \boldsymbol{\beta} + \epsilon_{Mi}, \\ \qquad \epsilon_{Mi} \overset{iid}{\sim} N(0, \sigma_M^2) \end{array}\right.$$

$$\text{Outcome model}\left\{\begin{array}{l} Y_i \;=\; h_Y(\mathbf{Z}_{Yi}, M_i) + \mathbf{C}_i^T \boldsymbol{\theta} + \epsilon_{Yi}, \\ \qquad \epsilon_{Yi} \overset{iid}{\sim} N(0, \sigma_Y^2) \end{array}\right.$$

$$\text{Total effect model}\left\{\begin{array}{l} Y_i \;=\; h_{TE}(\mathbf{Z}_{Yi}) + \mathbf{C}_i^T \boldsymbol{\gamma} + \epsilon_{TEi}, \\ \qquad \epsilon_{TEi} \overset{iid}{\sim} N(0, \sigma_{TE}^2) \end{array}\right.$$

Notation

Indices	$i = 1, \ldots, n$	subjects
	$\ell = 1, \ldots, L$	mixture components
Data	$\mathbf{Y} = (Y_1, \ldots, Y_n)^T$	normally distributed outcome variable
	$\mathbf{C}$	covariate matrix with rows $\mathbf{c}_i^T$
	$\mathbf{Z}$	general exposure matrix with rows $\mathbf{z}_i^T = (z_{i1}, \ldots, z_{iL})$
	M	normally distributed mediator variable
	$\mathbf{A}$	continuous exposure mixture matrix
	$\mathbf{E}_M$	continuous effect modifiers of A-M
	$\mathbf{E}_Y$	continuous effect modifiers of A-M-Y
	$\mathbf{Z}_M = (\mathbf{A}, \mathbf{E}_M)$	matrix for kernel in the mediator model
	$\mathbf{Z}_Y = (\mathbf{A}, \mathbf{E}_Y)$	matrix for kernel in the total effect model
	$(\mathbf{Z}_Y, M)$	matrix for kernel in the outcome model
Parameters	$\mathbf{h} = (h_1, \ldots, h_n)^T$	subject-specific health effects $h_i = h(\mathbf{z}_i)$
	$\mathbf{K}_{\mathbf{Z},\mathbf{r}}$	$n \times n$ kernel matrix for variable selection with (i,j)-element $\exp\left\{-\sum_{\ell=1}^L r_\ell (z_{i\ell} - z_{j\ell})^2\right\}$
	$\mathbf{r} = (r_1, \ldots, r_L)^T$	augmented variables in kernel matrix for variable selection, which controls smoothness of $h(\cdot)$
	$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_L)^T$	inclusion indicators for mixture components

Section A.2: Simulation details

For our simulation (Section 3), we used various linear and nonlinear functions to generate data. Graphical depictions of these functions are summarized in Figure 1 of the main text. The specific functions we used are the following:

$$
\begin{aligned}
h_{1,lin}(a_1) &= a_1 \\
h_{1,quad}(a_1) &= -\frac{2a_1^2}{3} + 3 \\
h_{2,lin}(a_1, a_2) &= \frac{a_1}{2} + \frac{a_2}{2} + \frac{a_1 a_2}{12} - \frac{3}{4} \\
h_{2,quad}(a_1, a_2) &= -\frac{(a_1 - 1)^2}{4} - \frac{(a_2 - 1)^2}{4} - \frac{(a_1 + 3)(a_2 + 3)}{3} + 5
\end{aligned}
$$

Section B: Algorithm to estimate CDEs using BKMR-CMA

1. Fit BKMR outcome model (9).

2. For each MCMC iteration, $j = 1, \ldots, J$:

   (a) Estimate the average outcome value for the mean level of covariates at the specific mediator value of interest for $\tilde{\mathbf{z}} = \begin{pmatrix} \mathbf{z}_Y & \mathbf{z}_Y^* \end{pmatrix}^T$ from (9) (i.e. estimate $Y_{\mathbf{z}^*m}$ and $Y_{\mathbf{z}m}$ for each MCMC iteration).

$$
\begin{aligned}
\mathbf{Y}_{\tilde{\mathbf{z}}m}^{(j)}(\bar{\mathbf{c}}) &= \mathrm{E}^{(j)}(Y|\mathbf{Z}_Y = \tilde{\mathbf{z}}, M = m, \mathbf{C} = \bar{\mathbf{c}}) \\
&= h_Y^{(j)}(\mathbf{Z}_Y = \tilde{\mathbf{z}}, M = m) + \bar{\mathbf{c}}^T \boldsymbol{\theta}^{(j)}
\end{aligned}
$$

   (b) Obtain the j^{th} posterior sample of the CDE for a change of exposure from $\mathbf{a}^*$ to $\mathbf{a}$ intervening to fix the mediator at m and effect modifiers $\mathbf{E}_Y$ at $\mathbf{e}_Y$ by:

$$
\mathrm{CDE}(m)^{(j)} = Y_{\mathbf{z}m}^{(j)}(\bar{\mathbf{c}}) - Y_{\mathbf{z}^*m}^{(j)}(\bar{\mathbf{c}}).
$$

3. Estimate the CDE(m) and its 95% credible interval conditional for level of the effect modifiers and marginally according to the confounders as the posterior mean and posterior 2.5^{th} and 97.5^{th} percentiles from these posterior samples.

Only two no unmeasured confounding assumptions are required for the CDE to have a causal interpretation: $Y_{\mathbf{a}m} \amalg \mathbf{A}|\mathbf{C}$, $Y_{\mathbf{a}m} \amalg M|\mathbf{C}, \mathbf{A}$. Namely, there are no unmeasured exposure-outcome confounders and there are no unmeasured mediator-outcome confounders.

Section C: Formulas to estimate causal mediation effects when the exposure is a mixture

Consider the following linear regression models for the mediator and outcome:

$$\mathrm{E}[M] = \beta_0 + \boldsymbol{\beta}_1^T \mathbf{A} + \boldsymbol{\beta}_2^T \mathbf{C}, \tag{C.1}$$

$$\mathrm{E}[Y] = \theta_0 + \boldsymbol{\theta}_1^T \mathbf{A} + \theta_2 M + \boldsymbol{\theta}_3^T \mathbf{A} M + \boldsymbol{\theta}_4^T \mathbf{C}, \tag{C.2}$$

where $\mathbf{A} = (A_1, \ldots, A_L)^T$ is a exposure mixture of L components, $\boldsymbol{\beta}_1 = (\beta_{11}, \ldots \beta_{1L})^T$, $\boldsymbol{\theta}_1 = (\theta_{11}, \ldots \theta_{1L})^T$, and $\boldsymbol{\theta}_3 = (\theta_{31}, \ldots \theta_{3L})^T$.

$\mathrm{E}_Y(Y_{\mathbf{a}m}|\mathbf{C} = \mathbf{c})$ represents the expected outcome value had everyone been exposed to level $\mathbf{a}$ and had their mediator been set to level m, fixing covariates to level $\mathbf{c}$. $\mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\mathbf{C} = \mathbf{c})$ represents the expected outcome value had everyone been exposed to level $\mathbf{a}$ and had their mediator been set to the level it would have taken if exposure is set to $\mathbf{a}^*$, fixing covariates to level $\mathbf{c}$. Then, considering models (C.1) and (C.2), and assuming (i) $Y_{\mathbf{a}m} \amalg \mathbf{A}|\mathbf{C}$, (ii) $Y_{\mathbf{a}m} \amalg M|\mathbf{C}, \mathbf{A}$, (iii) $M_{\mathbf{a}} \amalg \mathbf{A}|\mathbf{C}$, and (iv) $Y_{\mathbf{a}m} \amalg M_{\mathbf{a}^*}|\mathbf{C}$, we can estimate these effects as:

$$
\begin{aligned}
\mathrm{E}_Y(Y_{\mathbf{a}m}|\mathbf{C}) &\overset{(i)-(ii)}{=} \mathrm{E}_Y(Y|\mathbf{A}=\mathbf{a}, M=m, \mathbf{C}=\mathbf{c}) \qquad \text{by consistency} \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \theta_2 m + \boldsymbol{\theta}_3^T\mathbf{a}m + \boldsymbol{\theta}_4^T\mathbf{c} \\
\mathrm{E}_Y(Y_{\mathbf{a}^*m}|\mathbf{C}) &\overset{(i)-(ii)}{=} \mathrm{E}_Y(Y|\mathbf{A}=\mathbf{a}^*, M=m, \mathbf{C}=\mathbf{c}) \qquad \text{by consistency} \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a}^* + \theta_2 m + \boldsymbol{\theta}_3^T\mathbf{a}^*m + \boldsymbol{\theta}_4^T\mathbf{c}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\mathbf{C}) &= \int_m \mathrm{E}_Y(Y_{\mathbf{a}m}|M_{\mathbf{a}^*}=m, \mathbf{C}=\mathbf{c})dP_{M_{\mathbf{a}^*}}(m|\mathbf{C}=\mathbf{c}) \\
&\overset{(iii)}{=} \int_m \mathrm{E}_Y(Y_{\mathbf{a}m}|M_{\mathbf{a}^*}=m, \mathbf{C}=\mathbf{c})dP_{M_{\mathbf{a}^*}}(m|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \\
&\overset{(iv)}{=} \int_m \mathrm{E}_Y(Y_{\mathbf{a}m}|\mathbf{C}=\mathbf{c})dP_M(m|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \qquad \text{by consistency} \\
&\overset{(i)-(ii)}{=} \int_m \mathrm{E}_Y(Y|\mathbf{A}=\mathbf{a}, M=m, \mathbf{C}=\mathbf{c})dP_M(m|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \qquad \text{by consistency}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_m \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \theta_2 m + \boldsymbol{\theta}_3^T\mathbf{a}m + \boldsymbol{\theta}_4^T\mathbf{c}\; dP_M(m|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\mathbf{c} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\int_m m\; dP_M(m|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\mathbf{c} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\mathrm{E}_M(M|\mathbf{A}=\mathbf{a}^*, \mathbf{C}=\mathbf{c}) \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\mathbf{c} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\mathbf{c}\right]
\end{aligned}
$$

By similar logic,

$$
\begin{aligned}
\mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}}}|\mathbf{C}) &= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\mathbf{c} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a} + \boldsymbol{\beta}_2^T\mathbf{c}\right] \\
\mathrm{E}_Y(Y_{\mathbf{a}^*M_{\mathbf{a}^*}}|\mathbf{C}) &= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a}^* + \boldsymbol{\theta}_4^T\mathbf{c} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}^*\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\mathbf{c}\right]
\end{aligned}
$$

Thus,

$$
\begin{aligned}
CDE(m) &= \mathrm{E}_Y(Y|\mathbf{A}=\mathbf{a}, M=m, \mathbf{C}=\mathbf{c}) - \mathrm{E}_Y(Y|\mathbf{A}=\mathbf{a}^*, M=m, \mathbf{C}=\mathbf{c}) \\
&= \left(\boldsymbol{\theta}_1^T + \boldsymbol{\theta}_3^T m\right)(\mathbf{a}-\mathbf{a}^*) \\
NDE &= \int_{\mathbf{c}} \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\mathbf{C}) - \mathrm{E}_Y(Y_{\mathbf{a}^*M_{\mathbf{a}^*}}|\mathbf{C}) dP_{\mathbf{C}}(\mathbf{c}) \\
&\approx \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\bar{\mathbf{C}}) - \mathrm{E}_Y(Y_{\mathbf{a}^*M_{\mathbf{a}^*}}|\bar{\mathbf{C}}) \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\bar{\mathbf{c}} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\bar{\mathbf{c}}\right] - \\
&\quad \left(\theta_0 + \boldsymbol{\theta}_1^T\mathbf{a}^* + \boldsymbol{\theta}_4^T\bar{\mathbf{c}} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}^*\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\bar{\mathbf{c}}\right]\right) \\
&= \boldsymbol{\theta}_1^T(\mathbf{a}-\mathbf{a}^*) + \boldsymbol{\theta}_3^T(\mathbf{a}-\mathbf{a}^*)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\bar{\mathbf{c}}\right] \\
NIE &= \int_{\mathbf{c}} \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}}}|\mathbf{C}) - \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\mathbf{C}) dP_{\mathbf{C}}(\mathbf{c}) \\
&\approx \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}}}|\bar{\mathbf{C}}) - \mathrm{E}_Y(Y_{\mathbf{a}M_{\mathbf{a}^*}}|\bar{\mathbf{C}}) \\
&= \theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\bar{\mathbf{c}} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a} + \boldsymbol{\beta}_2^T\bar{\mathbf{c}}\right] - \\
&\quad \left(\theta_0 + \boldsymbol{\theta}_1^T\mathbf{a} + \boldsymbol{\theta}_4^T\bar{\mathbf{c}} + \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\beta_0 + \boldsymbol{\beta}_1^T\mathbf{a}^* + \boldsymbol{\beta}_2^T\bar{\mathbf{c}}\right]\right) \\
&= \left(\theta_2 + \boldsymbol{\theta}_3^T\mathbf{a}\right)\left[\boldsymbol{\beta}_1^T(\mathbf{a}-\mathbf{a}^*)\right]
\end{aligned}
$$

When considering traditional approaches to model the outcome, we do not include exposure-mediator interactions in (C.2). We therefore model the outcome as:

$$
\mathrm{E}[Y] = \gamma_0 + \boldsymbol{\gamma}_1^T\mathbf{A} + \gamma_2 M + \boldsymbol{\gamma}_3^T\mathbf{C}. \tag{C.3}
$$

We estimate the traditional mediation effects for an exposure mixture as:

$$
\begin{aligned}
NDE &= \boldsymbol{\gamma}_1^T(\mathbf{a}-\mathbf{a}^*), \\
NIE &= \theta_2\boldsymbol{\beta}_1^T(\mathbf{a}-\mathbf{a}^*).
\end{aligned}
$$

For both the linear and traditional methods, we model the total effect (TE) by:

$$\mathrm{E}[Y] \;=\; \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{A} + \boldsymbol{\alpha}_2^T \mathbf{C}, \tag{C.4}$$

and estimate the TE as: $\boldsymbol{\alpha}^T(\mathbf{a} - \mathbf{a}^*)$.

Section D: Supplementary Figures

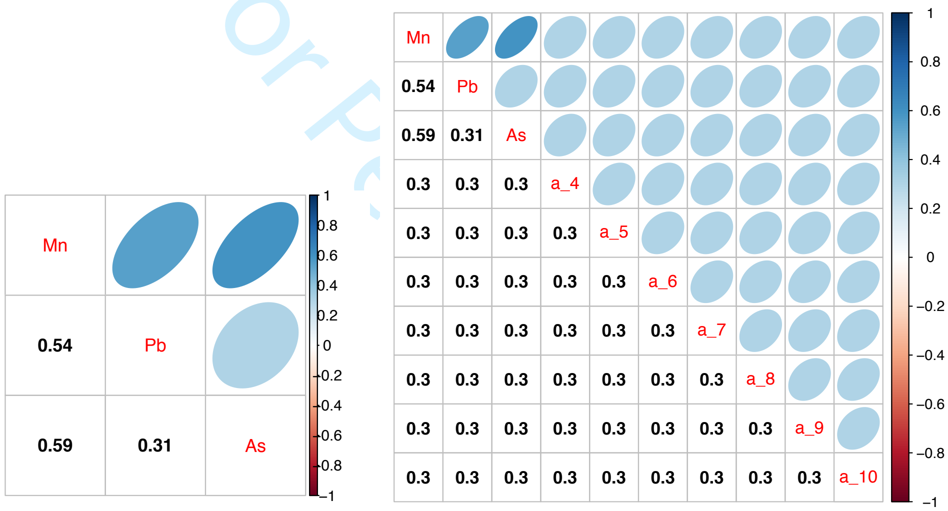

Figure D.1: Covariance structure $\boldsymbol{\Sigma}$ considered in our simulation when $L = 3$ and $L = 10$. The covariance for manganese (Mn), arsenic (As), and lead (Pb) from Bangladesh after log transformation and standardization.

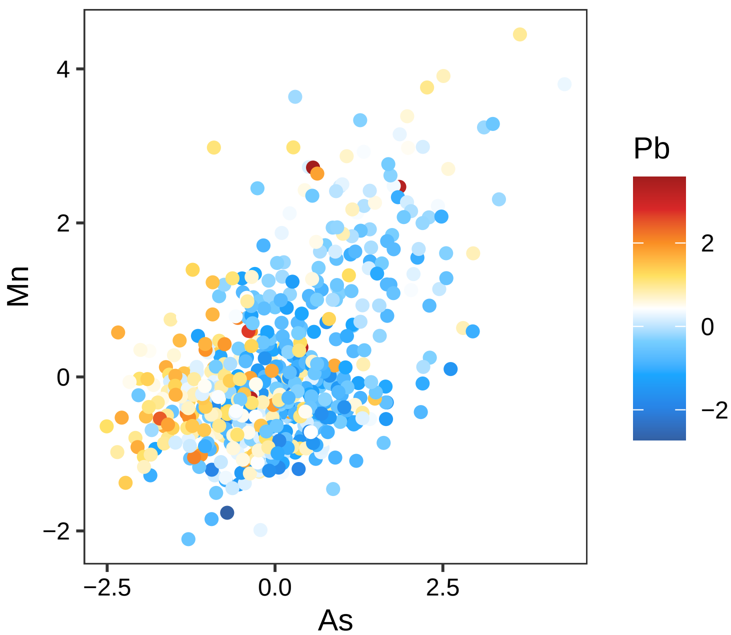

Figure D.2: Three dimensional scatter plot of natural logarithm transformed and standardized metal concentrations observed in our Bangladeshi cohort. The 25^{th} percentiles of the natural logarithm transformed and standardized metals are $As_{.25} = -0.61$, $Mn_{.25} = -0.64$, $Pb_{.25} = -0.80$, which corresponds to the raw levels of $As_{.25} = 0.56\mu$g/dL, $Mn_{.25} = 4.72\mu$g/dL, $Pb_{.25} = 1.15\mu$g/dL. The 75^{th} percentiles of the natural logarithm transformed and standardized of metals are $As_{.75} = 0.55$, $Mn_{.75} = 0.34$, $Pb_{.75} = 0.74$, which corresponds to the raw levels of $As_{.75} = 1.58\mu$g/dL, $Mn_{.75} = 17.80\mu$g/dL, $Pb_{.75} = 2.42\mu$g/dL.

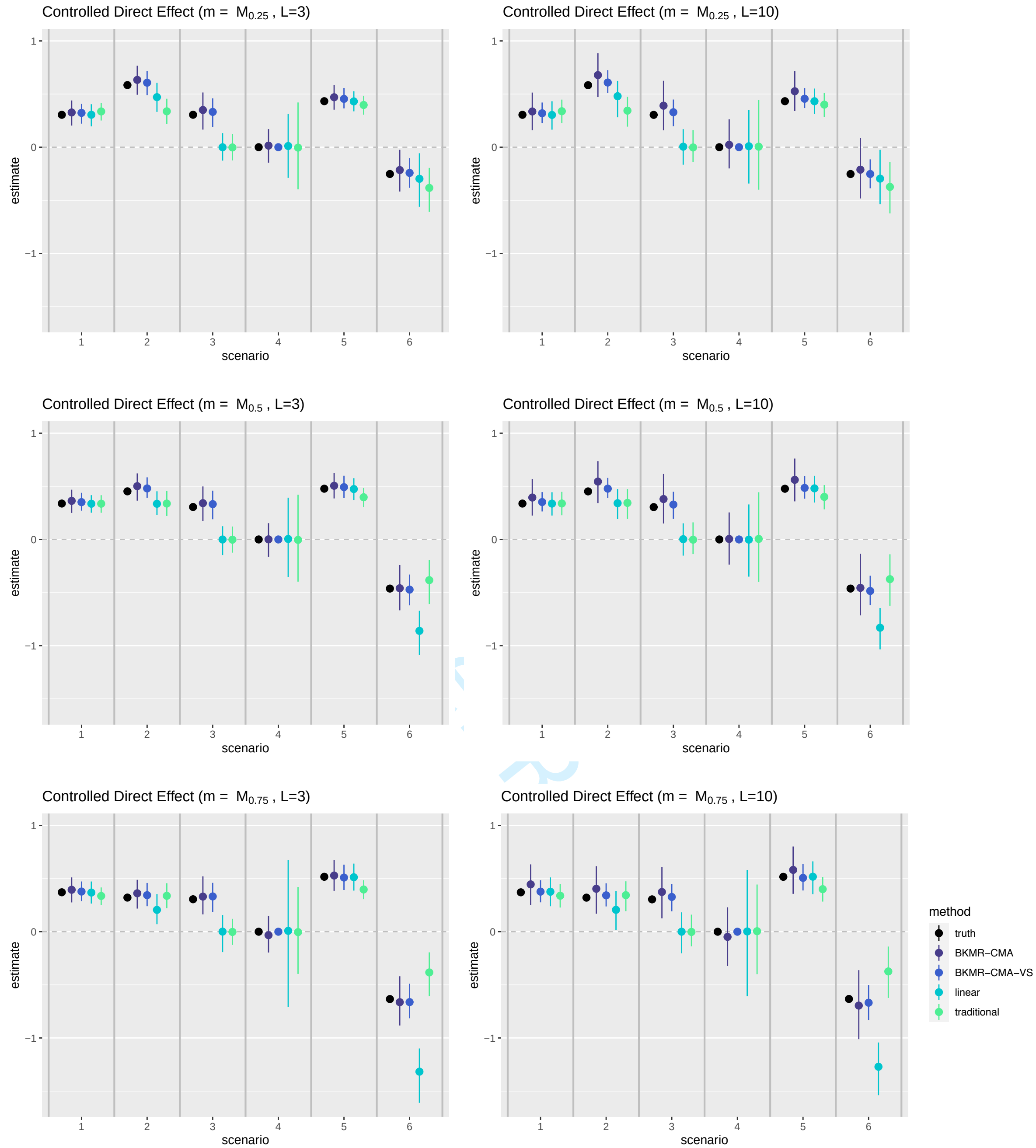

Figure D.3: Empirical median and 2.5^{th} and 97.5^{th} percentiles calculated from the estimates of the CDEs across the 500 simulation datasets using our proposed BKMR-CMA and BKMR-CMA-VS approaches, the linear method, and the traditional method under each simulation scenario. The CDEs presented are for when the mediator is fixed to its 25^{th}, 50^{th}, and 75^{th} percentiles in the true underlying dataset for each scenario. The truth for each mediation effect and scenario are depicted as black dots. The specific data generation functions used for each simulation scenario are defined in Table 1 of the main text. Results are show for six different data generation scenarios and when the number of mixture components is three and ten.

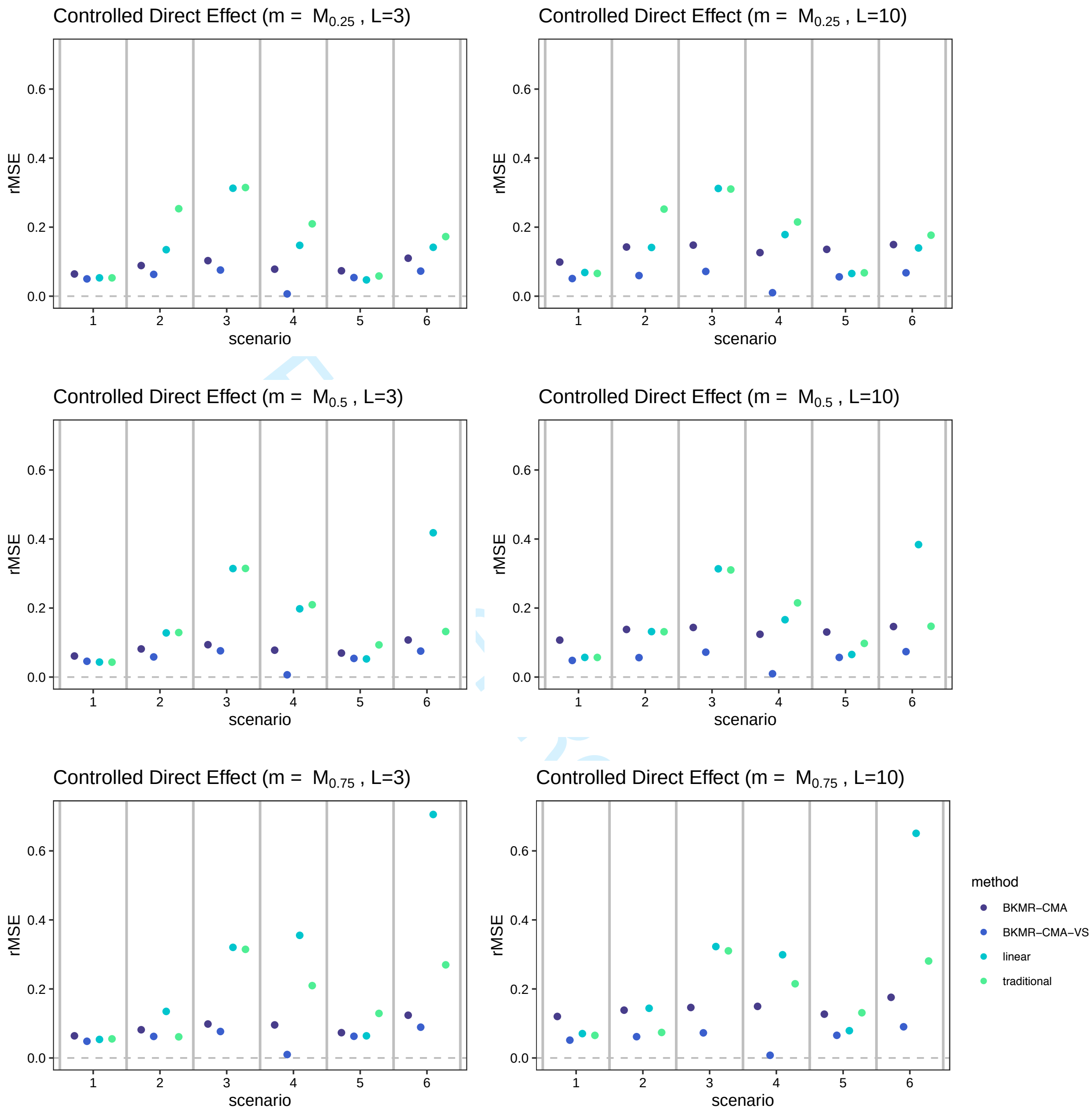

Figure D.4: A comparison of the root mean square error (rMSE) from our simulation when the CDEs are estimated by our BKMR-CMA and BKMR-CMA-VS approaches, the linear method, and the traditional method. The CDEs presented are for when the mediator is fixed to its 25^{th}, 50^{th}, and 75^{th} percentiles in the true underlying dataset for each scenario. Results are show for six different data generation scenarios and when the number of mixture components is three and ten. The specific data generation functions used for each simulation scenario are defined in Table 1 of the main text.

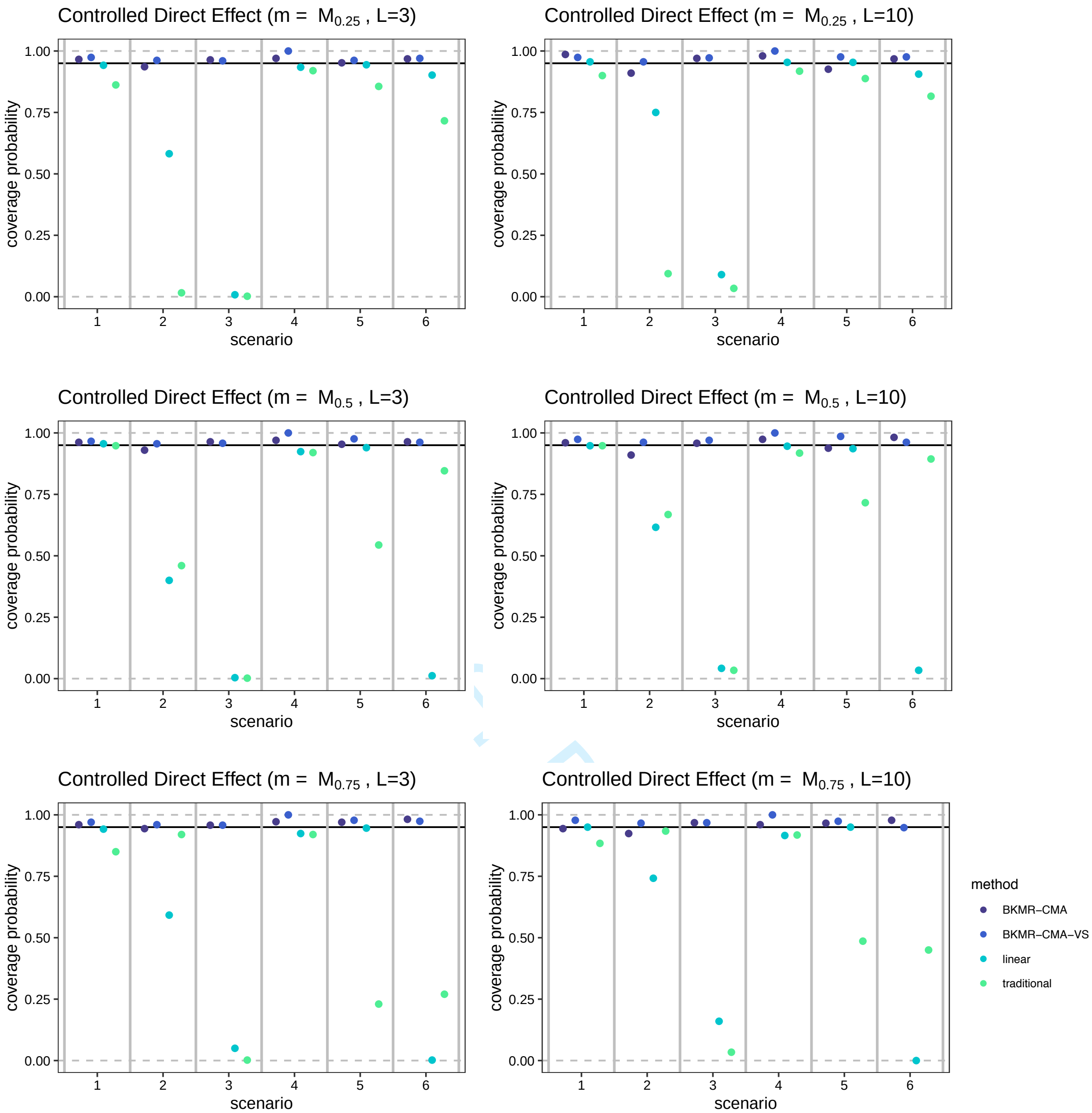

Figure D.5: A comparison of the coverage probability from our simulation when the CDEs are estimated by our BKMR-CMA and BKMR-CMA-VS approaches, the linear method, and the traditional method. The coverage probability is defined as the proportion of the estimates and 95% credible or confidence intervals in the 500 simulation datasets that contain the truth for each effect. The black line represents a coverage probability of 0.95. The CDEs presented are for when the mediator is fixed to its 25^{th}, 50^{th}, and 75^{th} percentiles in the true underlying dataset for each scenario. Results are show for six different data generation scenarios and when the number of mixture components is three and ten. The specific data generation functions used for each simulation scenario are defined in Table 1 of the main text.

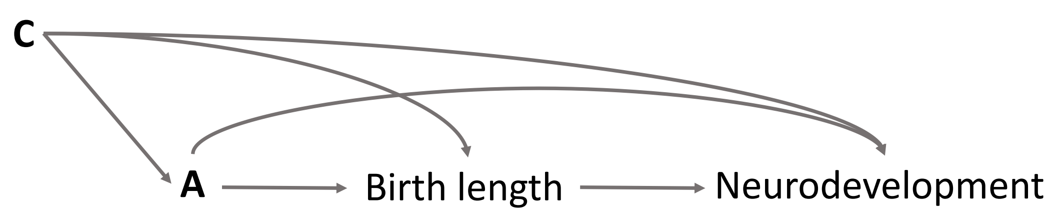

Figure D.6: Directed acyclic graph for the direct effect of a vector of correlated exposures ($\mathbf{A}$ = As, Mn, Pb) on children's neurodevelopment and for the indirect effect of *in utero* co-exposure to As, Mn, and Pb ($\mathbf{A}$) on children's neurodevelopment through birth length, where $\mathbf{C}$ denotes a vector of confounders.