Biostatistics (2017), **0**, 0, *pp.* **??**–16
doi:10.1093/biostatistics/LKMRSupplementRevised

Lagged kernel machine regression for identifying time windows of susceptibility to exposures of complex metal mixtures

Supplementary materials

SHELLEY H. LIU[1,*], JENNIFER F. BOBB[2], KYU HA LEE[3],

CHRIS GENNINGS[1], BIRGIT CLAUS HENN[4], DAVID BELLINGER[5],

CHRISTINE AUSTIN[1], LOURDES SCHNAAS[6], MARTHA TELLEZ-ROJO[7],

HOWARD HU[8], ROBERT O. WRIGHT[1], MANISH ARORA[1], BRENT A. COULL[5]

[1]*Icahn School of Medicine at Mount Sinai, New York, NY, U.S.A.*

[2]*Kaiser Permanente Washington Health Research Institute, Seattle, WA, U.S.A.*

[3]*The Forsyth Institute, Boston, MA, U.S.A.*

[4]*Boston University School of Public Health, Boston, MA, U.S.A.*

[5]*Harvard T.H. Chan School of Public Health, Boston, MA, U.S.A.*

[6]*National Institute of Perinatology, Mexico*

[7]*National Institute of Public Health, Mexico*

[8]*University of Toronto Dalla Lana School of Public Health, Canada*

shl159@mail.harvard.edu

[*]To whom correspondence should be addressed.

Section A: Form of covariance matrix Σ_h^{-1}

When G_t is an identity matrix, Σ_h^{-1} is a tridiagonal matrix. Otherwise, Σ_h^{-1} is a block diagonal matrix with blocks of size n, and two off-diagonals of length *n(T-1).*

In order to write out the form of Σ_h, we use the identities:

$$\frac{\lambda_1}{2\sigma}e^{\frac{-\lambda_1||\boldsymbol{h}_t||_{G_t}}{\sigma}} = \int_0^\infty (\frac{1}{2\pi\sigma^2\tau_t^2})^{1/2} exp(-\frac{||\boldsymbol{h}_t||_{G_t}}{2\sigma^2\tau_t^2}) \frac{(\frac{\lambda_1^2}{2})^{\frac{N+1}{2}}(\tau_t^2)^{\frac{N+1}{2}-1}}{\Gamma(\frac{N+1}{2})} e^{-\lambda_1^2\tau_t^2/2} d\tau_t^2$$

$$\frac{\lambda_2}{2\sigma}e^{\frac{-\lambda_2|h_{t+1,n}-h_{t,n}|}{\sigma}} = \int_0^\infty (\frac{1}{2\pi\sigma^2\omega_t^2})^{1/2} exp(-\frac{(h_{t+1,n}-h_{t,n})^2}{2\sigma^2\omega_t^2}) \frac{\lambda_2^2}{2} e^{-\lambda_2^2\omega_t^2/2} d\omega_t^2$$

Focusing on the exponents in the two identities that involve h,

$$\sum_{t=1}^{T} \frac{||\boldsymbol{h}_t||_{G_t}}{2\sigma^2\tau_t^2} + \sum_{n=1}^{N}\sum_{t=1}^{T-1} \frac{(h_{t+1,n}-h_{t,n})^2}{2\sigma^2\omega_t^2}$$

We see that $\boldsymbol{h}|\tau_1^2, ..., \tau_T^2, \omega_1^2, ..., \omega_{T-1}^2, \sigma^2$ is normally distributed with mean $\mathbf{0}$ and covariance matrix $\sigma^2\Sigma_h$, such that:

$$
\Sigma_h^{-1} = \begin{bmatrix}
\frac{G_{1,11}}{\tau_1^2} + \frac{1}{\omega_1^2} & \frac{G_{1,12}}{\tau_1^2} & \cdots & \frac{G_{1,1n}}{\tau_1^2} & \frac{-1}{\omega_1^2} & 0 & \cdots & 0 & \cdots \\
\frac{G_{1,12}}{\tau_1^2} & \frac{G_{1,22}}{\tau_1^2} + \frac{1}{\omega_1^2} & \cdots & \frac{G_{1,2n}}{\tau_1^2} & 0 & \frac{-1}{\omega_1^2} & 0 & \cdots & \cdots \\
\cdots & \cdots & \ddots & \cdots & \cdots & \ddots & \ddots & \cdots & \cdots \\
\frac{G_{1,1n}}{\tau_1^2} & \frac{G_{1,2n}}{\tau_1^2} & \cdots & \frac{G_{1,nn}}{\tau_1^2} + \frac{1}{\omega_1^2} & 0 & 0 & 0 & \frac{-1}{\omega_1^2} & \cdots \\
\frac{-1}{\omega_1^2} & 0 & \cdots & 0 & \frac{G_{2,11}}{\tau_2^2} + \frac{1}{\omega_1^2} + \frac{1}{\omega_2^2} & \frac{G_{2,12}}{\tau_2^2} & \cdots & \frac{G_{2,1n}}{\tau_2^2} & \ddots \\
0 & \frac{-1}{\omega_1^2} & \ddots & 0 & \frac{G_{2,12}}{\tau_2^2} & \frac{G_{2,22}}{\tau_2^2} + \frac{1}{\omega_1^2} + \frac{1}{\omega_2^2} & \cdots & \frac{G_{2,2n}}{\tau_2^2} & \cdots \\
\cdots & \cdots & \ddots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & \frac{-1}{\omega_1^2} & \frac{G_{2,1n}}{\tau_2^2} & \frac{G_{2,2n}}{\tau_2^2} & \cdots & \frac{G_{2,nn}}{\tau_2^2} + \frac{1}{\omega_1^2} + \frac{1}{\omega_2^2} & \cdots \\
\cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots
\end{bmatrix}
$$

Supplementary figure S1. Form of covariance matrix Σ_h^{-1}

Section B: Prior specification, MCMC sampler and full conditional distributions for lkmr

1. Prior specification

To complete the model specification, we define prior distributions for the regression parameters $\boldsymbol{\beta}$ and σ^2. As suggested by Park and Casella (2008), we take the prior for σ^2 to be inverse gamma,

$$\pi(\sigma^2) = \frac{\gamma^a}{\Gamma(a)}(\sigma^2)^{-a-1}e^{\frac{-\gamma}{\sigma^2}}, \sigma^2 > 0, a > 0, \gamma > 0 \tag{1.1}$$

and $\pi(\boldsymbol{\beta}) \propto \mathbf{1}$. As suggested by Kyung *and others* (2010), we use a gamma prior on λ_1^2 and λ_2^2. Specifically, priors are of the form

$$\pi(\lambda^2) = \frac{\delta^r}{\Gamma(r)}(\lambda^2)^{r-1}e^{-\delta\lambda^2}, \lambda^2 > 0, r > 0, \delta > 0. \tag{1.2}$$

In the simulation study described in Section 3 of the manuscript, we evaluated a grid of r and δ values to identify hyperparameters that optimize root mean squared error (RMSE) for estimation of the exposure-response surface across a range of realistic exposure scenarios. Our simulation studies identified hyperparameter values for λ_1^2 of $r_1 = 60$ and $\delta_1 = 10$, and hyperparameters for λ_2^2 of $r_2 = 45$ and $\delta_2 = 10$, as values that performed well under a range of simulation scenarios which considered differing numbers of toxicants and time windows, as well as differing exposure-response relationships.

2. MCMC sampler

In this section, we describe the Gibbs sampler implementation for the hierarchy of (2.7) - (2.10) of the main manuscript. For convenience, we denote $\sum_t \boldsymbol{h}_t$ as $\mathbf{W}\boldsymbol{h}$. The joint density is

$$f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{h}, \sigma^2, \boldsymbol{\tau}^2, \boldsymbol{\omega}^2) \propto \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[\frac{-1}{2\sigma^2}(\mathbf{y} - \mathbf{W}\boldsymbol{h} - \boldsymbol{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{W}\boldsymbol{h} - \boldsymbol{X}\boldsymbol{\beta})\right] \times$$

$$\prod_{t=1}^{T}\left(\frac{1}{2\pi\sigma^2\tau_t^2}\right)^{n/2}\exp\left(-\frac{\|\boldsymbol{h_t}\|^2_{\mathbf{G}_t}}{2\sigma^2\tau_t^2}\right)\frac{\left(\frac{\lambda_1^2}{2}\right)^{\frac{n+1}{2}}\left(\tau_t^2\right)^{\frac{n+1}{2}-1}}{\Gamma\left(\frac{n+1}{2}\right)}\exp\left(-\lambda_1^2\tau_t^2/2\right)\times$$
$$\prod_{t=1}^{T-1}\left(\frac{1}{2\pi\sigma^2\omega_t^2}\right)^{1/2}\exp\left[-\frac{\sum_{i=1}^{n}\left(h_{i,t+1}-h_{i,t}\right)^2}{2\sigma^2\omega_t^2}\right]\frac{\lambda_2^2}{2}\exp\left(-\lambda_2^2\omega_t^2/2\right)\times \tag{2.3}$$
$$(\sigma^2)^{-a-1}e^{\frac{-\gamma}{\sigma^2}}\times\prod_{r=1}^{2}(\lambda^2)^{r-1}e^{-\delta\lambda^2}.$$

The full conditional distribution of $\boldsymbol{h}$ is:

$$\boldsymbol{h}|\sigma^2,\boldsymbol{\tau}^2,\boldsymbol{\omega}^2,X,\boldsymbol{\beta},\mathbf{y}\sim N_{nT}\left[(\boldsymbol{W}^T\boldsymbol{W}+\Sigma_h^{-1})^{-1}\boldsymbol{W}^T\ (\mathbf{y}-\mathbf{X}\boldsymbol{\beta}),\sigma^2(\boldsymbol{W}^T\boldsymbol{W}+\Sigma_h^{-1})^{-1}\right]$$

The full conditional of σ^2 is:

$$\sigma^2|\boldsymbol{h},\boldsymbol{\tau}^2,\boldsymbol{\omega}^2,\boldsymbol{X},\boldsymbol{\beta},\mathbf{y}\sim \tag{2.4}$$
$$\textit{Inverse Gamma}\left[\frac{n(T+1)}{2}+a,\frac{(\mathbf{y}-\mathbf{W}\boldsymbol{h}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{W}\boldsymbol{h}-\mathbf{X}\boldsymbol{\beta})+\boldsymbol{h}^T\Sigma_h^{-1}\boldsymbol{h}}{2}+\gamma\right]$$

Also, the full conditional of β is:

$$\boldsymbol{\beta}|\boldsymbol{h},\sigma^2\sim N\left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{y}-\mathbf{W}\boldsymbol{h}),\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\right] \tag{2.5}$$

The full conditional distribution for $\boldsymbol{\tau}^2$ is:

$$\frac{1}{\tau_t^2}\sim \textit{Inverse Gaussian}\left[\sqrt{\frac{\sigma^2\lambda_1^2}{||\boldsymbol{h}_t||^2_{G_t}}},\lambda_1^2\right]\ \text{for } t=1,...,T.$$

The full conditional distribution for $\boldsymbol{\omega}^2$ is:

$$\frac{1}{\omega_t^2}\sim \textit{Inverse Gaussian}\left[\sqrt{\frac{\sigma^2\lambda_2^2}{\sum_{n=1}^{N}(h_{t+1,n}-h_{t,n})^2}},\lambda_2^2\right]\ \text{for } t=1,...,T-1.$$

Lastly, the full conditional distributions of λ_1^2 and λ_2^2, with priors $\Gamma(r_1,\delta_1)$ and $\Gamma(r_2,\delta_2)$, respectively, are:

$$\lambda_1^2\sim gamma\left[\frac{T(N+1)}{2}+r_1,\sum_{t=1}^{T}\frac{\tau_t^2}{2}+\delta_1\right]$$
$$\lambda_2^2\sim gamma\left[T-1+r_2,\sum_{t=1}^{T-1}\frac{\omega_t^2}{2}+\delta_2\right]$$

The Gibbs sampler is implemented to cyclically sample from the distributions of $\boldsymbol{h},\boldsymbol{\beta},\sigma^2$, $\boldsymbol{\tau^2}$, $\boldsymbol{\omega^2}$, λ_1^2, and λ_2^2 conditional on the current values of the other parameters. We note that for several parameters, $\boldsymbol{h}$, $\boldsymbol{\beta}$, $\boldsymbol{\tau}^2$, and $\boldsymbol{\omega}^2$, the Gibbs sampler is a block update.

Section C: Simulation results

For each scenario, we present the intercept, slope and R^2 of the regressions, as well as RMSE and coverage, over 100 simulations, each with a dataset of 100 subjects. Supplementary Table S1 considers a different exposure response function than used in Table 1 of the manuscript, in order to study the performance of LKMR when the shape of the exposure-response surface differs at different time windows. For this scenario, $h(\mathbf{z}) = 2 * z_{2,1}^2 + 2 * z_{2,3}^2 - 2 * z_{3,2}$, for $z_{t,m}$, $t = 1, ..., 4$ time windows and $m = 1, ..., 5$ toxicants, and an auto-correlation of 0.5 among toxicants. We see that for all time windows, LKMR has a lower RMSE and a higher R^2 than BKMR, suggesting a better estimation of the time-specific mixture effects. Furthermore, LKMR better estimates the mixture effects at time window 2 as compared with JKBKMR - the slope under LKMR is 0.81 compared with 0.39 for JKBKMR, and the RMSE is 1.09 under LKMR compared with 2.47 for JKBKMR. We also consider three other simulation scenarios which assume the same shape of the exposure-response surface as Table 1 of the manuscript, but for a larger number of toxicants and/or time windows.

We first assess the performance of LKMR for a larger number of toxicants. Supplementary Table S2 considers a scenario of 10 toxicants and 4 time windows. In this scenario, the RMSE under LKMR is consistently lower for all time windows than for BKMR or JKBKMR, with the except of time window 1. Furthermore, the R^2 is consistently higher across all time windows for LKMR as compared with BKMR or JKBKMR.

We next assess the performance of LKMR for a larger number of time windows. Supplementary Table S3 considers a simulation scenario of 10 time windows and 5 toxicants. Here, R^2 is consistently higher, coverage is consistently higher, and RMSE is consistently lower for all time windows except for time window 1 when comparing LKMR against BKMR or JKBKMR.

To complete the simulations, we finally assess the performance of LKMR for a larger number of toxicants and time windows. The results presented in Supplementary Table S4, which simulates

exposure data of 10 toxicants and 10 time windows, show that compared with BKMR, LKMR has a lower RMSE as well as higher coverage across all time windows. Furthermore, LKMR has consistently higher slope, R^2 and coverage than JKBKMR. We see that across all methods, when a larger number of time windows and/or larger number of metals is studied, while maintaining the same sample size, performance generally decreases. However, LKMR is still better able to estimate the time-specific exposure-response surface, even if its effects are sometimes attenuated.

Supplementary table S1. Regression of $\widehat{\boldsymbol{h}}$ on h under a different exposure-response simulation scenario for $M = 5$ toxicants and $T = 4$ time windows. $h(\mathbf{z}) = 2 * z_{2,1}^2 + 2 * z_{2,3}^2 - 2 * z_{3,2}$, for $z_{t,m}$, $t = 1, ..., 4$ time windows and $m = 1, ..., 5$ toxicants. Performance of estimated $h_t(\mathbf{z}_i)$ across 100 simulated datasets, each with N=100. AR-1 denotes the autocorrelation in simulated data. RMSE denotes the root mean squared error of the $\hat{h}$ as compared to h. Coverage denotes the proportion of times that the true h falls within in the posterior credible interval of each time point. At time windows 1 and 4, there is no effect; thus, slope and R^2 are not applicable to the regression of $\widehat{\boldsymbol{h}}$ on $\boldsymbol{h}$. BKMR refers to applying BKMR to exposure data from each time point separately. Joint Kernel refers to applying BKMR simultaneously to data from all time points.

h function	Time window	Intercept	Slope	R^2	RMSE	Coverage
LKMR	1	0.01	N/A	N/A	0.86	0.97
	2	-0.02	0.81	0.94	1.09	0.90
AR-1 = 0.5	3	0.02	0.72	0.78	0.95	0.92
	4	0.01	N/A	N/A	0.78	0.98
BKMR	1	0.05	N/A	N/A	1.16	0.95
	2	0.02	0.95	0.87	1.33	0.78
AR-1 = 0.5	3	0.06	0.87	0.73	1.20	0.94
	4	0.05	N/A	N/A	1.10	0.94
Joint Kernel	1	0.04	N/A	N/A	0.85	0.99
	2	0.13	0.39	0.71	2.47	0.81
AR-1 = 0.5	3	-0.03	0.86	0.83	0.90	0.99
	4	0.03	N/A	N/A	0.87	0.99

Supplementary table S2. Regression of $\widehat{\boldsymbol{h}}$ on h under an exposure-response simulation scenario for $M = 10$ toxicants and $T = 4$ time windows. We assumed there is no effect of exposure to the mixture at time $t = 1$, and a gradual increase in the effect over time, by defining $h_t(\mathbf{z}) = \alpha_t h(\mathbf{z})$, $\mathbf{z} = (z_1, ..., z_{10})^T$, where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = 2 * (0, 0.5, 0.8, 1)$ and $h(\mathbf{z}) = z_1^2 - z_2^2 + 0.5 z_1 z_2 + z_1 + z_2$. Performance of estimated $h_t(\mathbf{z}_i)$ across 100 simulated datasets, each with N=100. AR-1 denotes the autocorrelation in simulated data. RMSE denotes the root mean squared error of the $\hat{h}$ as compared to h. Coverage denotes the proportion of times that the true h falls within in the posterior credible interval of each time point. At time window 1, there is no effect; thus, slope and R^2 are not applicable to the regression of $\widehat{\boldsymbol{h}}$ on $\boldsymbol{h}$. BKMR refers to applying BKMR to exposure data from each time point separately. JKBKMR refers to applying BKMR simultaneously to data from all time points.

h function	Time window	Intercept	Slope	R^2	RMSE	Coverage
LKMR	1	0.05	N/A	N/A	1.74	1.00
	2	0.02	0.64	0.59	1.61	1.00
AR-1 = 0.5	3	-0.03	0.64	0.78	1.99	0.99
	4	0.05	0.72	0.82	2.17	0.99
BKMR	1	0.20	N/A	N/A	2.20	0.97
	2	0.16	1.15	0.50	3.08	0.94
AR-1 = 0.5	3	0.00	1.30	0.74	3.41	0.91
	4	0.24	1.10	0.76	3.16	0.92
JKBKMR	1	0.14	N/A	N/A	1.51	1.00
	2	0.15	0.49	0.42	2.11	0.99
AR-1 = 0.5	3	0.10	0.49	0.58	2.74	0.96
	4	0.15	0.47	0.60	3.29	0.91

Supplementary table S3. Regression of $\widehat{\boldsymbol{h}}$ on h under an exposure-response simulation scenario for $M = 5$ toxicants and $T = 10$ time windows. We assumed there is no effect of exposure to the mixture at time $t = 1$, and a gradual increase in the effect over time, by defining $h_t(\mathbf{z}) = \alpha_t h(\mathbf{z})$, $\mathbf{z} = (z_1, ..., z_5)^T$, where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = 2 * (0, 0.5, 0.8, 1)$ and $h(\mathbf{z}) = z_1^2 - z_2^2 + 0.5 z_1 z_2 + z_1 + z_2$. Performance of estimated $h_t(\mathbf{z}_i)$ across 100 simulated datasets, each with N = 100. AR-1 denotes the autocorrelation in simulated data. RMSE denotes the root mean squared error of the $\hat{h}$ as compared to h. Coverage denotes the proportion of times that the true h falls within in the posterior credible interval of each time point. At time windows 1, 5, 6, 7, 8, 9, 10, there is no effect; thus, slope and R^2 are not applicable to the regression of $\widehat{\boldsymbol{h}}$ on $\boldsymbol{h}$. BKMR refers to applying BKMR to exposure data from each time point separately. JKBKMR refers to applying BKMR simultaneously to data from all time points.

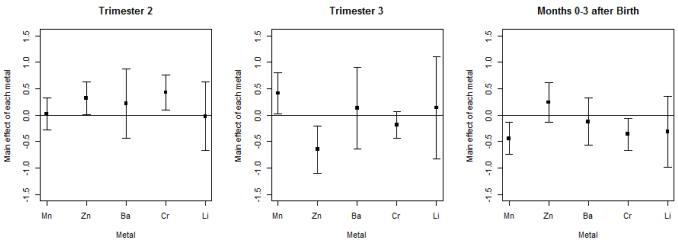
h function	Time window	Intercept	Slope	R^2	RMSE	Coverage
LKMR	1	0.05	N/A	N/A	1.21	1.00
	2	0.02	0.72	0.79	1.20	1.00
AR-1 = 0.5	3	0.07	0.76	0.91	1.45	1.00
	4	-0.04	0.66	0.92	1.98	0.97
	5	0.05	N/A	N/A	1.16	1.00
	6	0.02	N/A	N/A	0.83	1.00
	7	-0.01	N/A	N/A	0.80	1.00
	8	0.01	N/A	N/A	0.82	1.00
	9	0.01	N/A	N/A	0.79	1.00
	10	-0.02	N/A	N/A	0.99	1.00
BKMR	1	0.48	N/A	N/A	2.07	0.93
	2	0.33	1.34	0.65	2.85	0.90
AR-1 = 0.5	3	0.26	1.48	0.85	3.22	0.82
	4	0.17	1.22	0.87	2.73	0.87
	5	0.48	N/A	N/A	2.73	0.89
	6	0.47	N/A	N/A	1.75	0.95
	7	0.47	N/A	N/A	1.63	0.97
	8	0.47	N/A	N/A	1.55	0.98
	9	0.46	N/A	N/A	1.55	0.98
	10	0.45	N/A	N/A	1.47	0.97
JKBKMR	1	0.09	N/A	N/A	1.07	1.00
	2	0.08	0.40	0.51	1.98	0.98
AR-1 = 0.5	3	0.09	0.39	0.55	2.93	0.91
	4	0.09	0.33	0.54	3.78	0.85
	5	0.11	N/A	N/A	1.11	1.00
	6	0.08	N/A	N/A	1.04	1.00
	7	0.09	N/A	N/A	1.02	1.00
	8	0.09	N/A	N/A	1.01	1.00
	9	0.07	N/A	N/A	1.01	1.00
	10	0.08	N/A	N/A	1.02	1.00

Supplementary table S4. Regression of $\widehat{\boldsymbol{h}}$ on h under an exposure-response simulation scenario for $M = 10$ toxicants and $T = 10$ time windows. We assumed there is no effect of exposure to the mixture at time $t = 1$, and a gradual increase in the effect over time, by defining $h_t(\mathbf{z}) = \alpha_t h(\mathbf{z})$, $\mathbf{z} = (z_1, ..., z_5)^T$, where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = 4 * (0, 0.5, 0.8, 1)$ and $h(\mathbf{z}) = z_1^2 - z_2^2 + 0.5 z_1 z_2 + z_1 + z_2$. Performance of estimated $h_t(\mathbf{z}_i)$ across 100 simulated datasets, each with N=100. AR-1 denotes the autocorrelation in simulated data. RMSE denotes the root mean squared error of the $\hat{h}$ as compared to h. Coverage denotes the proportion of times that the true h falls within in the posterior credible interval of each time point. At time windows 1, 5, 6, 7, 8, 9, 10, there is no effect; thus, slope and R^2 are not applicable to the regression of $\widehat{\boldsymbol{h}}$ on $\boldsymbol{h}$. BKMR refers to applying BKMR to exposure data from each time point separately. JKBKMR refers to applying BKMR simultaneously to data from all time points.

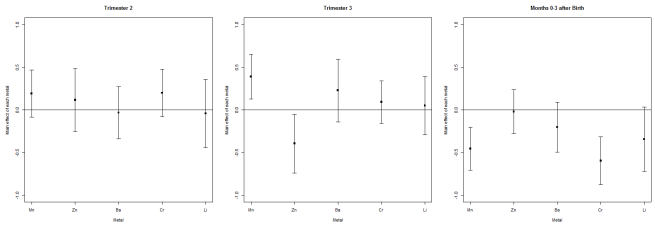
h function	Time window	Intercept	Slope	R^2	RMSE	Coverage
LKMR	1	0.13	N/A	N/A	3.02	1.00
	2	0.02	0.60	0.72	2.61	1.00
AR-1 = 0.8	3	-0.01	0.46	0.82	4.27	0.99
	4	-0.02	0.36	0.81	6.36	0.95
	5	-0.04	N/A	N/A	2.79	1.00
	6	-0.02	N/A	N/A	2.08	1.00
	7	-0.06	N/A	N/A	1.80	1.00
	8	-0.03	N/A	N/A	1.73	1.00
	9	-0.04	N/A	N/A	1.75	1.00
	10	-0.07	N/A	N/A	2.15	1.00
BKMR	1	-0.40	N/A	N/A	9.96	0.84
	2	-0.67	2.26	0.71	9.81	0.85
AR-1 = 0.8	3	0.01	2.04	0.88	10.02	0.78
	4	0.17	1.52	0.85	8.28	0.87
	5	-0.42	N/A	N/A	11.51	0.81
	6	-0.40	N/A	N/A	8.46	0.90
	7	-0.42	N/A	N/A	5.88	0.93
	8	-0.40	N/A	N/A	4.67	0.98
	9	-0.37	N/A	N/A	4.60	0.98
	10	-0.41	N/A	N/A	5.16	0.97
JKBKMR	1	0.00	N/A	N/A	1.85	1.00
	2	-0.09	0.30	0.47	3.98	0.96
AR-1 = 0.8	3	-0.04	0.23	0.47	6.27	0.87
	4	0.02	0.18	0.46	8.26	0.80
	5	-0.01	N/A	N/A	1.73	1.00
	6	0.00	N/A	N/A	1.48	1.00
	7	-0.03	N/A	N/A	1.40	1.00
	8	-0.02	N/A	N/A	1.27	1.00
	9	-0.06	N/A	N/A	1.34	1.00
	10	0.03	N/A	N/A	1.41	1.00

Section D: Application of bkmr and joint kernel bkmr to element data

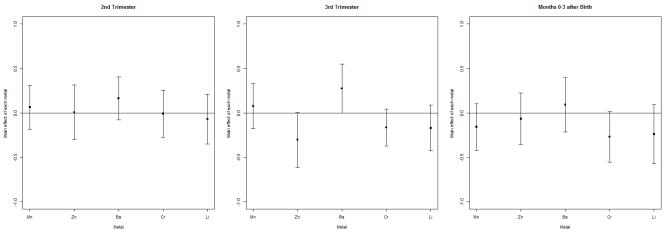
Linear regression applied to ELEMENT data



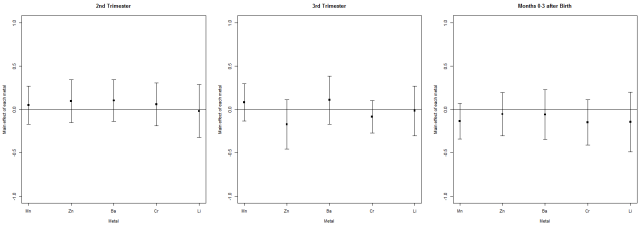
LKMR applied to ELEMENT data



BKMR applied to ELEMENT data

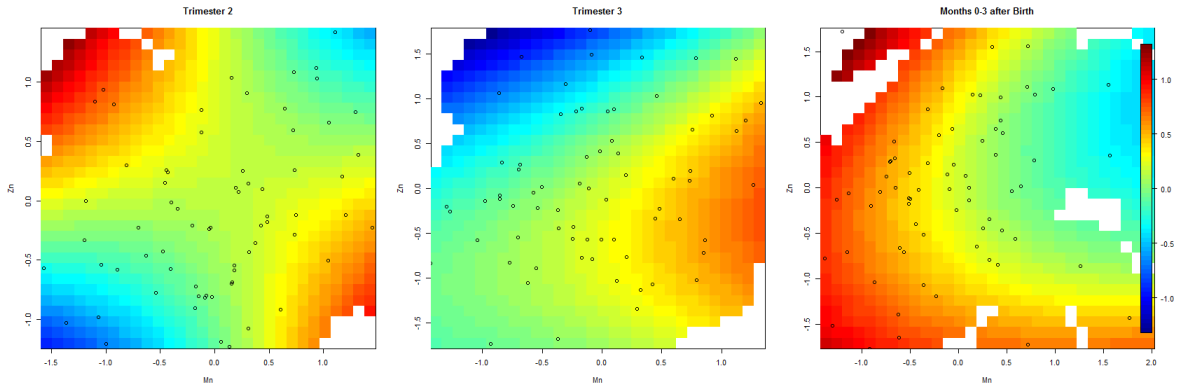


Joint kernel applied to ELEMENT data

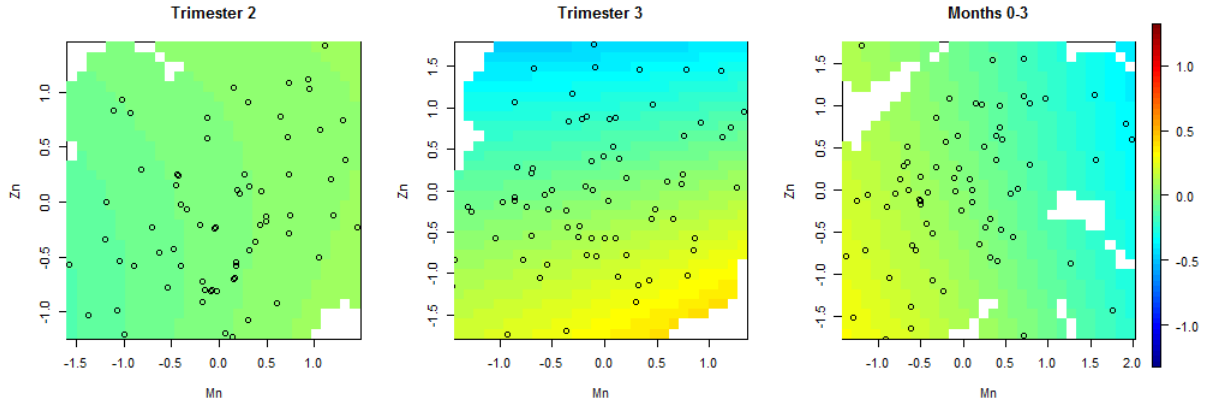


Supplementary figure S2. BKMR and joint kernel (JKBKMR) estimated main effect of each metal at three critical windows for ELEMENT data. LKMR plot is provided as a reference. Plot of the estimated relative importance of each metal, as quantified by the difference in the mean response at the 75th percentile versus the 25th percentile of a given metal exposure, while holding all other metal exposures constant at their median values.

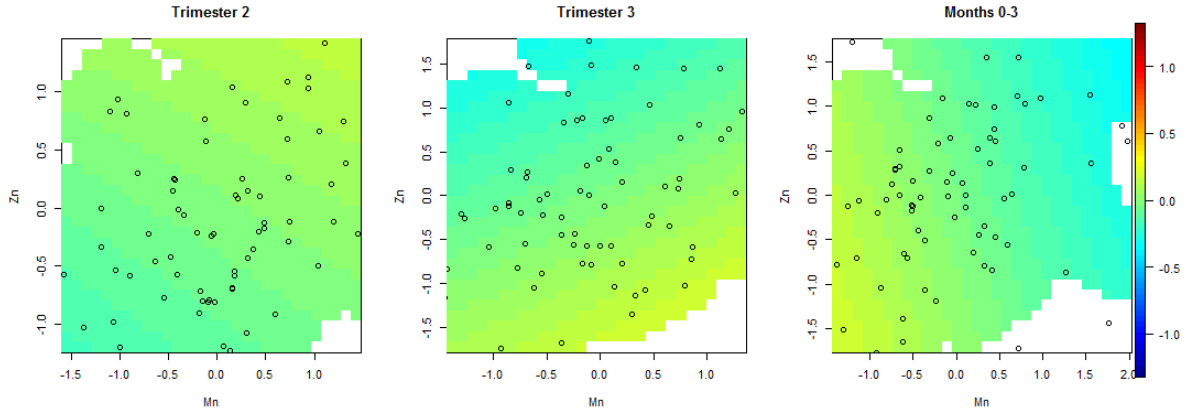
LKMR applied to ELEMENT data



BKMR applied to ELEMENT data

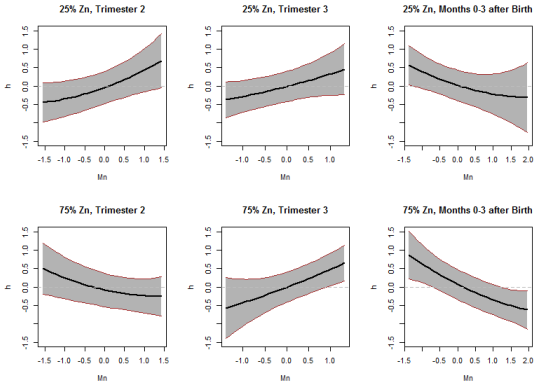


Joint kernel applied to ELEMENT data

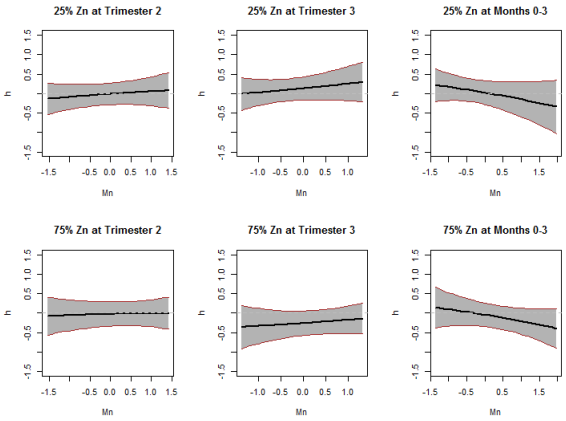


Supplementary figure S3. BKMR and joint kernel (JKBKMR) estimated time-specific exposure response functions applied to ELEMENT data. LKMR plot is provided as a reference. Plot of the estimated posterior mean of the exposure-response surface for Mn and Zn, at the median of Ba, Cr, and Li.

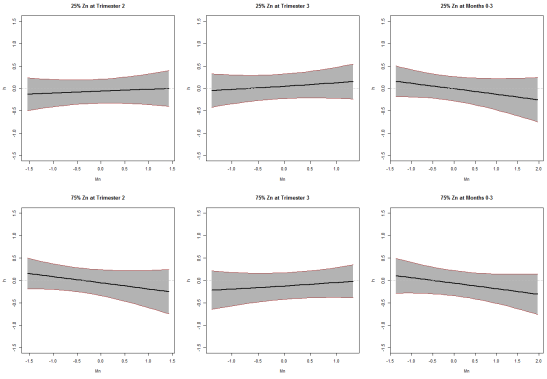
LKMR applied to ELEMENT data



BKMR applied to ELEMENT data



Joint kernel applied to ELEMENT data



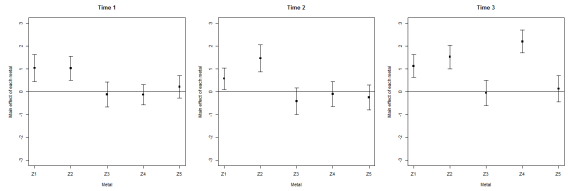
Supplementary figure S4. BKMR and joint kernel (JKBKMR) estimated time-specific exposure-response functions for Mn at low and high Zn levels applied to ELEMENT data. LKMR plot is provided as a reference. Plot of the cross-section of the estimated exposure-response surface for Mn, at the 25th (top panel) and 75th (bottom panel) of Zn exposure, holding Ba, Cr, and Li constant at median exposures.

Section E: Application of lkmr to a simulated case study dataset

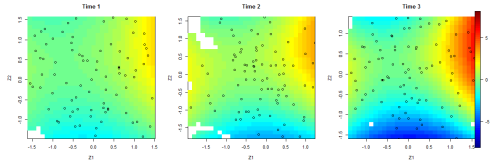
We describe the dataset and graphical results for a simulated case study. The simulated case study considered a five-toxicant scenario: three toxicants (out of five) exert a gradual non-additive, non-linear effect over three time windows that are representative of pregnancy and early life. We used the following model: $y_i = \mathbf{x}_i^T\boldsymbol{\beta} + \sum_t h_t\left(\mathbf{z}_{i,t}\right) + e_i$, where $\mathbf{z}_{i,t} = \left(z_{1i,t}, z_{2i,t}, z_{3i,t}\right)^T$, $e_i \sim N(0,1)$, $\mathbf{x}_i = (x_{1i}, x_{2i})^T$ and $x_{1i} \sim N\left(10,1\right)$ and $x_{2i} \sim Bernoulli(0.5)$. We simulated auto-correlation within toxicant exposures $\mathbf{z}_m = (\mathbf{z}_{m,1}^T, \mathbf{z}_{m,2}^T, \mathbf{z}_{m,3}^T)^T$ for metal $m = 1,2,3,4,5$ across time, and correlation among toxicants, using the Kronecker product for the exposure correlation matrix. Auto-correlation (AR-1) within toxicants was set to be 0.5. The shape of the exposure-response function $h_t(\mathbf{z}_{i,t})$, which was assumed to be the same at each time point, was simulated as quadratic with two-way interactions. We assumed there is a gradual increase in the effect over time, by defining $h_t(\mathbf{z}) = \alpha_t h(\mathbf{z})$ for time windows $t = 1,2$, where $\alpha = (\alpha_1, \alpha_2) = (0.5, 1)$ and $h(\mathbf{z}) = z_1^2 - z_2^2 + 0.5z_1z_2 + z_1 + z_2$, and $h_3(\mathbf{z}) = 1.5 * (z_1^2 - z_2^2 + 0.5z_1z_2 + z_1 + z_2 + z_4)$ for time window $t = 3$. We used the quadratic kernel function, such that $K(\mathbf{z},\ \mathbf{z'}) = (\mathbf{z}^T\mathbf{z'} + 1)^2$.

The description and analysis code for the simulated case study can be found in the Github repository: https://github.com/shelleyhliu/LKMR-CaseStudy.

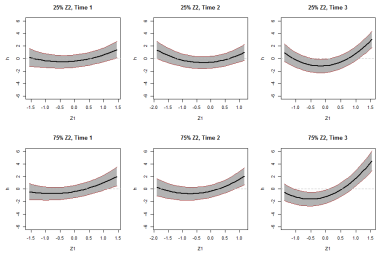
A. Relative importance of 5 simulated metals at 3 time windows



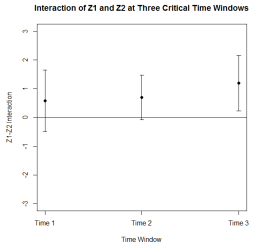
B. Heatmap of Z1-Z2 joint effects for 3 time windows



C. Cross-sectional plots of Z1 effects at low/high Z2 for 3 time points



D. Overall interaction between Z1-Z2 at 3 time points



Supplementary figure S5. LKMR estimated plots for a simulated case study dataset. A) Plot of the estimated relative importance of each metal, as quantified by the difference in the mean response at the 75th percentile versus the 25th percentile of a given metal exposure, while holding all other metal exposures constant at their median values. B) Plot of the estimated posterior mean of the exposure-response surface for Z1 and Z2, at the median of Z3, Z4, Z5. C) Plot of the cross-section of the estimated exposure-response surface for Z1, at the 25th (top panel) and 75th (bottom panel) of Z2 exposure, holding Z3, Z4, Z5 constant at median exposures. D) Plot of the estimated interaction effect between Z1 and Z2, holding Z3, Z4, Z5 constant at median exposures. First, we estimated the exposure-response effect for high (75th) versus low (25th) Z1 exposures, at high Z2 levels. Next, we estimated the exposure-response effect for high versus low Z1 exposures, at low Z2 levels. The difference between the two estimated exposure-response effects quantifies the Z1-Z2 interaction. All other metal exposures are held constant at their median values.

References

Kyung, M, Gill, J, Ghosh, M and Casella, G. (2010). Penalized regression, standard errors and Bayesian lassos. *Bayesian Analysis)* **5(2)**, 369–412.

Park, T and Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association* **103(482)**, 681–86.