Published in final edited form as:

Environ Sci Pollut Res Int. 2014 September ; 21(18): 10775–10786. doi:10.1007/s11356-014-3072-8.

Bayesian importance parameter modeling of misaligned predictors: soil metal measures related to residential history and intellectual disability in children

Georgiana Onicescu,

Department of Public Health Sciences, Medical University of South Carolina, Carolina, SC 29425, USA

Andrew B. Lawson,

Department of Public Health Sciences, Medical University of South Carolina, Carolina, SC 29425, USA

Suzanne McDermott,

Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA

C. Marjorie Aelion, and

School of Public Health and Health Sciences, University of Massachusetts Amherst, Amherst, MA 01003, USA

Department of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA

Bo Cai

Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA

Abstract

In this paper, we propose a novel spatial importance parameter hierarchical logistic regression modeling approach that includes measurement error from misalignment. We apply this model to study the relationship between the estimated concentration of soil metals at the residence of mothers and the development of intellectual disability (ID) in their children. The data consist of monthly computerized claims data about the prenatal experience of pregnant women living in nine areas within South Carolina and insured by Medicaid during January 1, 1996 and December 31, 2001 and the outcome of ID in their children during early childhood. We excluded mother-child pairs if the mother moved to an unknown location during pregnancy. We identified an association of the ID outcome with arsenic (As) and mercury (Hg) concentration in soil during pregnancy, controlling for infant sex, maternal race, mother's age, and gestational weeks at delivery. There is some indication that Hg has a slightly higher importance in the third and fourth months of

pregnancy, while As has a more uniform effect over all the months with a suggestion of a slight increase in risk in later months.

Keywords

Intellectual disabilities; Importance parameter; Measurement error; Soil metals; Bayesian analysis; Markov chain Monte Carlo

Introduction

Background

Intellectual disability (ID; previously classified as mental retardation) is a heterogeneous set of impairments characterized by significant limitations in both intellectual functioning and adaptive behavior. ID is the most common developmental disability, impacting approximately 6.5 million people in the USA. Approximately 545,000 children receive special education services for ID each year (National Dissemination Center for Children with Disabilities 2014). The average cost of caring for a child with ID could be ten times larger than the cost of caring for a child without any disability (Goldman and Koduru 2000). Among factors associated with ID are genetic conditions, social deprivation, toxicological exposures, injuries, and both prenatal and postnatal infections. However, in the great majority of the children, the cause still remains unknown (Goldman and Koduru 2000). While ID can result from a variety of causes, the toxicological exposures deserve special attention since they are preventable.

Humans routinely come in contact with chemical substances in the environment by inhalation, ingestion, or dermal contact with the soil, water, food, or other environmental media. The developing fetus is particularly sensitive and can be exposed to chemicals that cross the placenta. Like essential metals, most heavy metals can move rather freely across the human placenta. Therefore, the potential health effects of chemical exposure during pregnancy cannot be overlooked (Alberta Biomonitoring Program 2008). Because brain development starts at conception, it is important to understand mothers' risk factors during pregnancy that may pose a threat to the normal brain development of the fetus. Some toxicological/chemical factors linked to the development of ID during pregnancy are fetal alcohol syndrome and heavy metal poisoning. Chemicals have a complex distribution in the soil of rural and urban areas as a result of natural and anthropogenic sources (Davis et al. 2009). The main mechanism by which chemicals can cross the placenta and enter the fetus is through passive diffusion (Audus 1999).

This study was designed to test a novel approach to incorporate residential history. Our model describes the relationship of the binary intellectual disability (ID) outcome with the exposure to chemicals during the 9 months of pregnancy, estimating a common slope multiplied by importance parameter coefficients for the monthly chemical values. In this paper, we develop a methodological approach to assess mothers' exposure during pregnancy to chemicals found in the environment. The possibility of an exposure link between

concentrations of chemicals found in soils and pregnancy could have implications for prevention.

Maternal residence during pregnancy

Ignoring population mobility in studying the disease patterns can lead to bias of the risk estimates, with even a small amount of differential migration being able to bias the assessment of outcome and exposure relationships (Tong 2000). Including residential history in statistical models can provide insight on the location and the timing of environmental exposures.

Soil chemical composition

In this study, we considered a range of such chemicals thought to have possible importance in adverse perinatal health outcomes: Arsenic (As), Chromium (Cr), Mercury (Hg), Lead (Pb), Manganese (Mn), Barium (Ba), Copper (Cu), and Nickel (Ni).

Numerous reports document elevated soil metal concentrations in urban areas from industrial and transportation sources and elevated rural soil concentrations of metals from natural geologic sources, pesticides, and industrial facilities (Aelion et al. 2008; Davis et al. 2009, X Li et al. 2004). As is a chemical present in higher quantities in soils treated with pesticides (Walsh et al. 1977). Studies in animals suggest that ingested inorganic arsenic may produce developmental effects at high doses that also produce overt maternal toxicity (Centers for Disease Control and Prevention 2007a). High levels of As have been found to increase the probability of ID in children (Liu et al. 2010; McDermott et al. 2012, 2011; Vahter 2009). There is evidence of an association between As in soil and the combined outcome of ID and developmental delay (McDermott et al. 2011); however, there is no evidence of an association of As combined with other metals in soil and the outcome of ID (Liu et al. 2010). Cr is an element present in the soils due to both natural or from industrial releases, and different forms of Cr have varying degrees of toxicity. While some compounds are considered nutrients and required in small amounts for normal health functioning, others are more deleterious to health, especially when ingested in excessive doses (U.S. Department of Health And Human Services 2012). There are no studies regarding developmental effects in humans or animals after inhalation exposure to chromium or its compounds (Centers for Disease Control and Prevention 2012).

Hg is a naturally occurring metal found throughout the environment. Hg ingestion in high doses can affect several organ systems (Centers for Disease Control and Prevention 1999). Development of the fetus' central nervous system can be sensitive to methylmercury exposure. Methylmercury concentrations in blood serum of pregnant women have been found to vary by age and geographical location (Alberta Biomonitoring Program 2008). Prenatal exposure to Hg has been found to be associated with ID in children (McDermott et al. 2012; Trasande et al. 2006). Pb is a metal found in small amounts in the earth's crust. People are exposed to Pb via dust, old lead-based paints, drinking water, food, or drink. High soil concentration of Pb has been found to be associated with a higher probability of ID in children, and the Centers for Disease Control and Prevention now states that even low levels of Pb are a concern (Liu et al. 2010; McDermott et al. 2012). Mn is an essential

nutrient, with small amounts necessary for normal body functioning; however, in excess, it can be detrimental to health. Elevated hair Mn concentrations have been reported in children with learning disabilities (Collipp et al. 1983).

Barium is a metal which exists in nature only in ores containing mixtures of elements. The potential for Ba to induce reproductive and developmental effects has not been well investigated (Centers for Disease Control and Prevention 2007b). Copper is a metal that occurs naturally in rock, soil, water, sediment, and, at low levels, air, either as the free metal or associated with other elements in compounds that comprise various minerals. No studies were located regarding developmental effects of humans following oral exposure to copper (Centers for Disease Control and Prevention 2004). Nickel combined with other elements occurs naturally in the earth's crust. Human studies that examined whether nickel can harm the developing fetus are inconclusive (Centers for Disease Control and Prevention 2005).

Data

Health outcome

This research is motivated by a study of ID cases in children who were born to mothers who were pregnant, and insured by Medicaid, in South Carolina between 1996 and 2002 (Aelion et al. 2009a, b; Kim et al. 2009, 2010; Liu et al. 2010; McDermott et al. 2012, 2011; Zhen et al. 2008). In our analysis, we included only the first child for each mother. We deleted all mother-child pairs when a child was identified as having a genetic or other known cause of intellectual disability since the objective of this study was to study pregnancy-related inorganic metal exposures. Thus, cases of Down syndrome, Fragile X syndrome, Prader-Willi syndrome, Angelman syndrome, Williams syndrome, and fetal alcohol syndrome were taken out of the dataset since their cause was known. Medicaid is a health insurance program for individuals with low incomes in the USA, covering low-income parents, children, seniors, and people with disabilities.

The outcome assessment was based on two methods. First, we used the International Classification of Disease to identify children with an ICD9 code of 317–319. These three codes identify mild, moderate, severe, and other specified or unspecified mental retardation. The second approach used the identification of children in public school special education placement for educable mental disability (mild mental retardation), trainable mental disability (moderate mental retardation), or profound mental disability (severe or profound mental retardation). ID (previously termed mental retardation) can be caused by an insult that occurs during the perinatal period or after birth (Centers for Disease Control and Prevention & Intellectual Disability Fact Sheet 2014). There is a lag in the diagnosis of ID in children, the time of diagnosis depending on the presentation of signs and symptoms of known syndromes, and the observation of developmental delay in the developing child. Previous literature suggests the diagnosis of ID peaks around ages 9–11 years or at fifth grade (Pless 1994, S McDermott et al. 2007). In this study, the children were 6–11 years of age when our case identification for the ID diagnosis was carried out. We observed that child age had a nonlinear association with ID, and it started to level out at approximately age 7 years.

Our case definition had four levels: normal, mild mental retardation (ICD9 code 317 and/or placement in a public school for educable or mild MR), moderate/severe mental retardation (ICD9 318 and/or placement in a public school for trainable (moderate) or profound MR), and unspecified MR (ICD9 code 319). The objective of our study was to model the relationship between mental retardation (all severities) vs. the children without MR (typical school placement without an ICD9 code for 317–319). Our model is constructed for a binary outcome so we grouped the three levels of MR (mild, moderate/severe, and unspecified) and created a reference group (no MR).

We assessed the outcome using electronic administrative medical and school records that are merged with a unique individual identifier, and we allowed for 10 years of follow-up data for the children. During each month of pregnancy, residential addresses for mothers were recorded. Other available mother and child characteristics that were obtained from the birth certificate and maternal Medicaid record include the following: parity, infant sex, mother's alcohol consumption, mothers' smoking status, and mothers' race. We also linked census block group characteristics for median residential building age and density in the block group.

Sampling design and measurement

The study included nine areas where soil sampling could be carried out. These areas were selected because they contained a risk gradient for ID. In order to detect the areas of elevated risk, we used a local likelihood clustering methodology according to which the conditional logistic local likelihood formulation is derived by considering the probability of a location being a case given the joint distribution of cases and comparisons within the study window (Lawson 2006). The procedure for cluster analysis is described by Zhen et al. (Zhen et al. 2008). Each of the nine study areas was rectangular with an area of approximately 40 square miles. The areas were dispersed throughout South Carolina, in rural and urban locations. The original data included 9,402 mother-child pairs from the nine areas, which were previously described (Cai et al. 2012; McDermott et al. 2012; Zhen et al. 2008). The latitude and longitude of the four corners of the nine rectangular areas that contained a cluster of ID were identified, and a uniform grid was overlaid at locations 1.0–3.0 km apart. The coordinates for each area were mapped and Global Positioning System (GPS) latitudes and longitudes were taken at each sampling location with a handheld GPS device (Garmin Etrex, Olathe, KS) (Aelion et al. 2008). Soil was collected at 5-cm depths from each node (Aelion et al. 2009a, b); where nodes were inaccessible (e.g., on building locations or water bodies), soil samples were collected as close to the grid node as possible. The mean number of sampling sites in each study area is 115 with a standard deviation of 5. After sampling, soil was analyzed for the chemical values by an independent analytical laboratory (Pace Analytical, Huntersville, NC).

Data quality for residential data

We had the address of each woman for the months of pregnancy when she was enrolled in Medicaid. Since chemical values could not be measured within the private property domain of each individual residence, the concentrations of each metal were interpolated. Women were enrolled in Medicaid throughout pregnancy with a peak enrollment in month six. After

month 6, enrollments dropped off due to premature deliveries and maternal movement outside the nine areas for which soil sampling was conducted. The metal concentrations for each month of pregnancy were obtained as follows: if a mother lived at the same location as in month 6, for that month, we used the month 6 chemical value. If a mother moved in a certain month inside an area with known chemical values, we interpolated and obtained the chemical values at that location. If a mother moved less than 0.25 miles outside one of the nine study areas, we considered that the mother moved to a known address, and we used the chemical value for the closest location. Otherwise, if a mother moved to a location greater than 0.25 miles outside one of the nine study areas, we considered that the mother moved to an unknown location. If a mother moved to an unknown location but less than 0.5 miles from an area with known chemical values, we used the chemical values interpolated from the closest area. If a mother moved outside the original area but more than 0.5 miles from an area with known chemical values, the chemical value for that particular month was considered missing. The cut-offs were chosen after examination of the spatial correlation of the fields. The movement summary is described in Fig. 1.

The minimum distance from a particular latitude and longitude location to an area was determined by computing the minimum geo-distance between that particular location and each location on the borders of the area. The geo-distance between two locations was computed using the Vincenty distance formula (Vincenty 1975), taking into account the spherical shape of the Earth, with longitudes being closer to each other, the further the latitude is from the equator.

Data subsetting

Because of the complexity of the data structure, missing addresses, and movement of the subjects throughout the pregnancy, we classified each mother's status into one of the following categories: (1) subject did not move; (2) subject moved to a location with known chemical value; (3) subject moved to an address where the chemical value is unknown but inside the original area; (4) subject moved to a location with unknown chemical value outside the original area but less than half a mile from original areas; (5) subject moved to a location outside original area with unknown chemical value and more than half a mile from any original area; and (6) missing value for the latitude/longitude location coordinates.

The data used in the current analysis are a subset of the original data in which we excluded subjects who moved to an unknown location outside original areas and more than half a mile from any original area for which chemical values could not be obtained. Subjects with missing locations were included, and latitude and longitude values for the missing location were imputed using a last observation carried forward approach. If the first month location was missing, all the locations prior to the first non-missing location were imputed with the first non-missing latitude and longitude values. The data excluding subjects with missing covariates used for the adjusted models and including only the first pregnancy had a sample size of 6,889 subjects.

Statistical model

Various approaches have been used in the literature to model residential history (Cook et al. 2009; Hoffman et al. 2010; Kim et al. 2009, 2010; Manjourides and Pagano 2011; Van den Hooven et al. 2012; Vieira et al. 2002, 2005; Wheeler et al. 2012). Recently, with the advancement of geographical information systems, it has become easier to obtain point location data, resulting in an increased need for the development of statistical methods appropriate for this type of data. Unfortunately, most studies are only descriptive of the spatial distribution of disease since one cannot easily adjust for covariates or assess the risk of disease associated with a specific exposure. Existing non-Bayesian methods that allow the inclusion of covariates are generalized additive models (GAM) smoothing on location (Vieira et al. 2005) and universal Kriging, an extension of regular kriging allowing for the inclusion of spatially related covariates. An extension of the GAM model weighting each residential location by residential duration is proposed by Wheeler et al. (Wheeler et al. 2012).

Residential history data during the prenatal period and early childhood have been analyzed in the literature using one residential address in a certain period, such as the address on a birth certificate, (Hoffman et al. 2010) or separate monthly analysis during pregnancy (McDermott et al. 2011). Therefore, methodological approaches are needed to take into account the monthly movement of mothers during pregnancy.

Misalignment of soil chemical data

Since the soil sample sites are spatially misaligned to the residential addresses of pregnant women, we considered a range of options for the modeling of the misalignment. Due to confidentiality arrangements, we could not consider direct joint modeling of the soil chemicals and health outcomes (Banerjee et al. 2003; Gryparis et al. 2009; Lawson 2013; Paciorek 2012; Szpiro et al. 2011). Instead, we had to consider separate modeling of the interpolation of chemical measures. Let $(s_i, z(s_i))$ for $i=1, \dots, n$ denote the measurement of the random variable $(z(s))$ at the sampling location s_i . A simplified version of the model with one latent spatial process and nugget effect (zero) was considered, which can be expressed in a hierarchical scheme as:

$$\begin{aligned} \text{Level 1: } & z_i = F_i \alpha + G(s_i) \\ \text{Level 2: } & G(s) \sim N(0, \sigma_z^2 \Sigma(h, \phi)) \\ \text{Level 3: } & \text{prior}(\alpha, \sigma_z^2, \phi) \end{aligned}$$

In level 1, F represented a matrix of spatially referenced non-random variables at location $s_i=(x_{1i}, x_{2i})$ and α is a parameter vector. In the simple case considered here, we assumed a spatial linear trend for coordinates: $s_i=(x_{1i}, x_{2i})$, and, hence, the i^{th} row of F is $F_i=(1, x_{1i}, x_{2i})$.

In level 2, $G(s)$ is a stationary spatial Gaussian process (mean=0, variance= σ_z^2) and covariance function $\Sigma(h, \phi)$, where ϕ is the correlation parameter and h is the vector between two locations (i.e., $h_{lk}=\|s_l-s_k\|$). A range of parameterized covariances is available. From initial exploratory analysis, we found that semi-variograms of the soil chemicals were close

to exponential form, and so we adopted for simplicity an exponential correlation model in the analysis:

$$\Sigma(h_{lk}, \phi) = \exp(-h_{lk}/\phi).$$

The third level specifies the prior for the model parameters $(\alpha, \sigma_z^2, \phi)$, with a flat prior set for α and ϕ and the prior for σ_z^2 is σ_z^{-2} . We obtained posterior estimates of the model parameters and also the posterior mean levels of the chemicals, while also interpolating to the human health observation sites using the predictive distribution. Bayesian kriging was used for spatial interpolation of the chemicals and is described in more details elsewhere (Cai et al. 2011; Diggle and Ribeiro 2007). Sensitivity analysis was performed for the spatial correlation models used, and it was found that the kriged prediction distributions were comparable to metal internode concentration variation (Aelion et al. 2009a, b). The kriged estimates were validated using a leave-one-out cross validation (LOOCV) statistical method which compared the measured and LOOCV predicted values. Validation was also carried out using additional field sampling of soil metal concentrations at points between original sampling locations, which were compared to kriging prediction distributions. LOOCV results suggested that Bayesian kriging was a good predictor of metal concentrations (Aelion et al. 2009a, b).

Because of the misalignment of the addresses to the sampling sites, there is a measurement error problem with interpolation. One approach to this problem is to use a form of Berkson error (Banerjee et al. 2003; Carroll et al. 2006) where the interpolated value of a chemical is associated with a standard error estimated from the interpolation procedure (see “Spatial logistic model Description” for further discussion).

Spatial logistic model description

We assume that we observe a realization of a spatial point process with a binary outcome observed at each location defined by its latitude and longitude coordinates.

The locations are mothers' residential addresses during pregnancy with binary mark

$$y_i = \begin{cases} 1 & \text{if child has ID} \\ 0 & \text{otherwise.} \end{cases}$$

Mothers have residential addresses for each of 9 months during pregnancy. Denote these as $s_i = (s_{1i}, s_{2i}, s_{3i}, s_{4i}, s_{5i}, s_{6i}, s_{7i}, s_{8i}, s_{9i})^T$, where s_{li} are the geographical locations for mother-child pair i in month l , $s_{li} \in \mathbb{R}^2$, $i=1, \dots, n$; $l=1, \dots, 9$.

Although our environmental chemical covariates are measured longitudinally, the outcome is measured once after 10 years of follow-up of the children. Most ID is actually present at birth but the diagnosis is often made years later when a child cannot make expected progress in school. Clearly more cases of moderate, severe, and profound ID are diagnosed during the

preschool years. Since mild ID accounts for approximately 85 % of all ID cases, almost all cases are captured by the time children reach age 10 and are in fifth grade.

We assume a logistic spatial model at the first level of our hierarchical model whereby the binary ID outcome is related to the spatial covariates and effects. We term this model “an importance parameter logistic model”, and for the analyses, we adopted a Bayesian approach. Our models estimated a common overall slope, multiplied by importance coefficients for the chemical values over the nine months of pregnancy. We modeled the probability of having ID using a logit link function. Denote the true set of chemicals for the i^{th} site (individual) as a matrix with elements z_{kli}^* , $k=1, \dots, p; l=1, \dots, 9$ where p denotes the number of chemicals measured and l is the month of residence.

The model is specified as follows:

$$y_i \sim \text{Bern}(p_i)$$

$$\text{logit}(p_i | q_i, z^*(s_i)) = \beta_0 + \sum_{k=1}^p \beta_{1k} w_k^T z_{ki}^* + q_i^T \beta_2,$$

where $z_{ki}^* = (z_{k1i}^*, \dots, z_{k9i}^*)^T$ is the k th true latent soil chemical vector for the i th individual. $q_i = (q_1, \dots, q_m)^T$ is a vector of individual level covariates corresponding to the regression parameters $\beta_2 = (\beta_{21}, \dots, \beta_{2m})$, m being the number of individual level covariates. The vector $w_k = (w_{k1}, \dots, w_{k9})^T$ consists of the month specific importance parameters for the k^{th} chemical.

We denote z_{kli}^0 the interpolated (observed) chemical values at mothers' location s_{il} . In our study, the true chemical values $\{z_{kli}^*\}$ are not observed directly at mother's residence sites, but they are observed at fixed sample sites which are misaligned with mother's residences. The chemical values at mother's residences are here predicted using Bayesian Kriging. Therefore, our model must take into account the uncertainty resulting from this prediction. There are multiple ways in which measurement error could be taken into account in a statistical model (Carroll et al. 2006). As we must consider the interpolated values of chemicals as our observed data (due to confidentiality), we cannot consider a full joint modeling approach. Instead, we must consider a Berkson measurement error model, in which the observed covariate is assumed to have a prior normal distribution with mean the true covariate value and specified variance. Hence, we assume, for the vector of chemicals considered $\mathbf{G}_{li} = \log(\mathbf{z}_{li}^*)$ that:

$$G_{kli} \sim N\left(\log(z_{kli}^0), \tau_{\log z_{kli}^0}^{-1}\right)$$

$$\text{where, } \tau_{\log z_{kli}^0}^{-1} = \sigma_{z_{kli}^0}^2 / (z_{kli}^0)^2.$$

Here, $\sigma_{z_{kli}^0}^2$ is the estimated variance of the observed chemical based on the Bayesian kriged interpolation, and $\tau_{\log z_{kli}^0}^{-1}$ was obtained from the standard deviation of the observed values $\sigma_{z_{kli}^0}$ using the Delta method approximation, from the first order Taylor series expansion

around the mean: $\sigma_{\log Z^0_{kli}} = \frac{1}{z^0_{kli}} \sigma^0_{kli}$. Note that the variance of the observed chemical can vary with month as the mothers can move to different locations.

Other prior distributions

Our models are implemented in a Bayesian framework. Zero mean Gaussian prior distributions with variances $\sigma^2_{\beta_g}$ are assigned to the intercept and common slope:

$\beta_g \sim N(0, \sigma^2_{\beta_g})$; $g=0, 1$; $g = 0, 1$. The hyper-parameter standard deviation σ_{β_g} were assigned a uniform prior on the interval (0,C): $\sigma_{\beta_g} \sim U(0, C)$, $g = 0, 1$ (Gelman 2006). In examples, we examined a range of possible values for C, such as 2, 5, or 10, being mindful that this upper limit should not be informative. The month specific importance parameters $w_{kl1} = 1, \dots, 9$ were assigned the prior distribution:

$$R = \log \left\{ \frac{L(y|\theta') \Pi g_i(\theta')}{L(y|\theta) \Pi g_i(\theta_i)} \right\},$$

where p_k is the prior weight probability assumed to have value $C_0 = 0.5$.

Bayesian inference for estimation of parameters is based on the posterior distribution that is obtained by multiplying the likelihood and prior distributions of parameters.

Computational issues

Samples were generated from the posterior distribution using the Metropolis version of the Metropolis-Hastings (M-H) algorithm. Due to the large number of observations, the complexity of our model and the large number of iterations required to reach convergence, the run time of the program in available Bayesian software packages exceeded normal limits. Therefore, we implemented the M-H algorithm using the C++ software, which is a relatively fast intermediate level programming language. Using a CentOS Linux operating system with a Rocks Cluster installed, the run time was 3 days and 3 h for the final model and 2 days for the one chemical models adjusted for covariates. “Appendix 1: Metropolis-Hastings Algorithm” describes the M-H algorithm used.

Goodness of fit and convergence diagnostics

Each model was run for a varying number of iterations and discarded burn-in samples depending on convergence. For some models, thinning of the samples was performed to increase computational efficiency. In the Bayesian framework, model assessment and choices of the best fitting model can be performed using the DIC (Spiegelhalter et al. 2002), which is a Bayesian analog of the Akaike's information criteria (AIC). pD represents the effective number of parameters, which reflects the model complexity or degrees of freedom. Lower values of DIC indicate a better fit of the model. It is suggested that models with DIC within 1–2 units of the “best” deserve consideration and 3–7 have considerably less support (Tong 2000). Convergence diagnosis was performed using Geweke criteria implemented in the CODA package in R and which was based on a test of the equality of the means in the

first and last 25 % of the chain. The Geweke Z score is computed as the difference between the sample means divided by the estimated standard error. An absolute value of the Z score lower than 1.96 was considered to indicate convergence. Trace plots of the parameters were also used to graphically assess the convergence of the chain. The parameters had an acceptance probability of at least 20 %.

Results

Descriptive characteristics

Table 1 presents descriptive infant, maternal, and neighborhood characteristics of our dataset. Using simple logistic regression, when comparing the children with ID and children with no ID, those with ID were more likely to be male, have fewer weeks of gestation, and be small for gestational age. Maternal risk factors included a race other than non-Hispanic white, lower parity and alcohol use. Neighborhood characteristics were not statistically significantly associated with the outcome. Among the chemicals, the soil concentration for Pb, Hg, and Ba were significantly associated with ID.

Model comparison for importance parameter logistic models

Table 2 displays the DIC and the parameter degrees of freedom (pD). The lowest DIC was 2,923.52 for the full model including As and Hg chemicals and adjusted for infant sex, race, parity, mother's age, and gestational weeks, the DIC is 7.29 units smaller than the smallest DIC for the single chemical models. Since the smaller DIC was not close in value to any competing model, it was not necessary to use the parameters degrees of freedom (pD) as the secondary criteria to determine the best model.

Importance parameters logistic model results

A set of models was run for single chemical models adjusted for infant, maternal, and neighborhood characteristics, with continuous covariates which have been standardized for computational efficiency. The overall slope coefficient indicated an association for As and Hg.

The chemicals As and Hg were included in a full model, together with covariates that had an association in single chemical adjusted models: infant sex, race, parity, mother's age, and weeks of gestation.

The results for the full model are presented in Tables 3 and 4. The estimated common slope for As was 0.013 (0.00074, 0.038). The importance parameters were estimated to be 0.43, 0.45, 0.44, 0.46, 0.49, 0.47, 0.47, 0.48, and 0.47 for months 1 to 9 of pregnancy respectively, suggesting relatively uniform pregnancy effect over all the months of pregnancy, but with a suggestion of a slight increase in risk in later months. The common slope for Hg was 0.11 (0.021, 0.39). The estimated importance parameters were 0.49, 0.48, 0.53, 0.52, 0.40, 0.40, 0.37, 0.38, and 0.37 for the months 1 to 9 of pregnancy respectively. The importance parameters represent the probability that at a given month a particular chemical is in the model (i.e., that the effect of exposure in that month is non zero). All the credible intervals for these parameters fall in the range {0, 1} because they are based on sorted samples of 0 s

and 1 s due to the Bernoulli prior distribution assumption for the importance weights. There is some indication that Hg has slightly higher importance in the third and fourth months of pregnancy but the evidence is weak. The slope for the other chemicals did not show an association of the chemicals with the outcome. All the covariates included in the final model had an association with the outcome with the coefficients estimated at 0.69 (95 % CI=(0.48, 0.90)), 0.30 (95 % CI=(0.056, 0.56)), 0.19 (95 % CI=(0.092, 0.29)) and -0.30 (95 % CI=(-0.38, -0.22)) for infant sex, black race, mother's age, and weeks. Because of the standardization of the continuous variables mother's age and weeks, the interpretation of their coefficients is limited and these variables were used mainly for model adjustment.

Discussion

In this paper, we proposed a Bayesian importance parameter logistic regression model to identify the importance of certain predictors in their association with the outcome, including measurement error from misalignment. The model was applied to investigate the association between maternal exposure during pregnancy and intellectual disability in children.

The main finding of the study is that As and Hg are associated with ID, this being consistent with previous studies (McDermott et al. 2011). In our preliminary analyses we have fitted a frequentist multiple logistic regression model including all the chemicals, each averaged over all the months of pregnancy. We have fitted models both unadjusted and adjusted for infant sex, alcohol, black race, other race, smoking, building age of the house, density, parity, mother's age, and weeks. These models both showed an association between Hg exposure and ID, as well as borderline significance for As and Ba. Our model further suggests a more uniform effect of As with a slight increase in later months and higher importance of Hg exposure during the third and fourth months of pregnancy. This information is not available from a simpler logistic model approach.

We were concerned that the upper limits for the prior distributions of the standard deviations could be informative. However, examination of trace plots suggested that the standard deviations were not bounded by these limits. Also, in our analysis, we have adopted the simpler approach where the prior weight probability p_k is constant and equal to 0.5. A variety of forms can be assumed for this vector. A sophisticated variant could be to assume prior correlation between elements such as $\text{logit}(p_k) \sim N(\text{logit}(p_k^0), \Sigma)$ with p_k^0 being a null probability vector and correlation between months being specified by Σ . However, this may require prior knowledge for the specification of the null probability vector.

Although the use of C++ software increased the speed of running the models, due to the complexity of the modeling approach and the large dataset, the running time was still relatively long.

A limitation of our modeling approach is the fact that the model assumes linearity of the predictors and it does not consider their interactions. While a generalization of the model addressing these limitations is possible, it may limit interpretability and it may add additional complexities that may significantly increase computation time. Also, since the importance parameters can take only values of 0 and 1 in each step of the M-H algorithm,

the credible intervals around the estimated mean of the importance parameters are not very informative. In our study, we considered only the chemical presence in the soil at residential address, while mother's exposure can also be due to other sources, such as presence of chemicals in the atmosphere, ingestion of mercury-contaminated seafood, or food sourced from other locations.

As is the case with many epidemiological studies, there is the possibility of potential uncontrolled confounding such as other environmental pollutants or factors associated with the socioeconomic determinants of health. Since we only included Medicaid mother-child pairs, the variability in income was limited, but this did not eliminate all other elements of socioeconomic status.

The strength of our model is the incorporation of the importance parameters, which gives insight on the association between each month's chemical values and the outcome, as well as accounting for the kriged standard deviations associated with each chemical value.

Acknowledgments

Source of financial support: funding for this research was provided by the National Institutes of Health, R01 Grant No. ES012895-04A2.

Appendix 1: Metropolis-Hastings algorithm

Recent development in computer facilities made Markov chain Monte Carlo (MCMC) (Trasande et al. 2006) simulation techniques one of the main computational tools in Bayesian statistical inference. For this study, we implemented the Metropolis version of the Metropolis-Hastings algorithm (U.S. Department of Health And Human Services 2012), which is one of the most popular MCMC methods. The algorithm consists of the following steps repeated T times, where T is the number of iterations of the chain. Our parameter θ consists of the true chemical values $Z^*(s_{j,i})$, the regression coefficients and their standard deviations:

$\beta_0, \sigma_{\beta_0}, \beta_1, \sigma_{\beta_1}, (w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9), \beta_{21}, \sigma_{\beta_{21}}, \dots, \beta_{2m}, \dots, \sigma_{\beta_{2m}}, m$ being the number of individual level covariates.

In order to enhance computational efficiency block updates have been used for updating the true chemical values. Symmetric proposal distributions have been used for all the parameters.

Step 0 Assign starting values to θ .

For $t=1, \dots, T$:

Step 1 Propose new values θ' from symmetric proposal distributions ($h(\theta'|\theta)$). In this case, we have used Gaussian proposals centered on the previous value for continuous parameters on $(-\infty, \infty)$

Step 2 Calculate $\log(\alpha) = \min(0, R)$ then, as the symmetric proposals cancel in the ratio,

$$\mathbf{w}_k \sim \text{Bern}(p_k)$$

where L is the likelihood function and g is the prior distribution. The following prior distributions have been assigned:

$$\begin{aligned}\beta_l &\sim N(0, \sigma_{\beta_l}^2), l=0, 1 \\ \sigma_{\beta_l} &\sim U(0, 2), l=0, 1 \\ w_i &\sim \text{Bernoulli}(0.5), i=1, \dots, 9 \\ \log(Z^*(s_{j,i})) &\sim N(\log(Z^0(s_{j,i})), \sigma_{\log}^2(Z^0(s_{j,i}))), \\ j &= 1, \dots, 9, i=1, \dots, n\end{aligned}$$

Step 3 Update $\theta^{(t)} = \theta$ with probability α

Appendix 2: simulation

To test the ability of the model to estimate importance parameters when correlated with the outcome, we used simulated data with different model assumptions. The simulation was performed with two scenarios, one in which month six values were simulated from a different distribution compared to the rest of the months, with a higher association with the outcome (Model 1). The second scenario was there was no variation in the months related to the chemical concentrations, and they were all generated from the same distributions (Model 2). Each dataset had 500 observations. We included only a limited number of scenarios due to difficulty in obtaining convergence for other variety of distribution assumptions and parameters. In Model 1, the outcome was generated from a Bernoulli (0.5) distribution. The chemical values for each month except month six were generated from a lognormal distribution with log mean -8 and log standard deviation 0.1 to which, if the outcome was 1 a random lognormal with mean -8 and log standard deviation 0.1 was added for each individual in each month except month 6. In order to induce a higher association with the outcome, a random lognormal with mean -4.2 and log standard deviation 0.1 was added for month six. In Model 2, the outcome was generated from a Bernoulli (0.5) distribution. The chemical values for each month were generated from a lognormal distribution with log mean -8 and log standard deviation 0.0001 , to which a uniform $U(0, 0.0001)$ random number was added to each month values for each subject. The kriged standard deviation was fixed at 0.001 for both models. We considered a low standard deviation in order to have less impact on the estimated values. Model 1 was run for 20,000 iterations, with the first 5,000 iterations discarded as burn-in. Model convergence was checked using the Geweke criteria. The highest importance parameter was obtained for the month 6 values (0.99), with lower importance parameters for the rest of the months. For model 1, the credible intervals for the importance parameters were (0, 1) except for month 6 which was (1, 1). These results suggest a good capability of the model to identify the months that are correlated with the outcome. The mean intercept and 95 % credible interval was estimated to be 0.047 (-0.11 , 0.21), while the common mean slope and 95 % credible interval was 2.57 (2.54, 2.61). Model 2 was run for 20,000 iterations, 5,000 burn-in and a thinning of 3. Similar importance

parameters have been estimated for each month (Table 5). For Model 2 the credible intervals for the importance parameters were (0, 1).

References

- Aelion CM, Davis HT, McDermott S, Lawson AB. Metal concentrations in rural topsoil in South Carolina: potential for human health impact. *Sci Total Environ*. 2008; 402:149–156. [PubMed: 18538375]
- Aelion CM, Davis HT, Liu Y, Lawson AB, McDermott S. Validation of Bayesian kriging of arsenic, chromium, lead, and mercury surface soil concentrations based on internode sampling. *Environ Sci Technol*. 2009a; 43:4432–4438. [PubMed: 19603658]
- Aelion CM, Davis HT, McDermott S, Lawson AB. Soil metal concentrations and toxicity: associations with distances to industrial facilities and implications for human health. *Sci Total Environ*. 2009b; 407:2216–2223. [PubMed: 19155049]
- Alberta Biomonitoring Program. [07 March 2014] Chemicals in serum of pregnant women in Alberta.. Alberta Health and Wellness. 2008. <http://www.health.alberta.ca/documents/Chemical-Biomonitoring-2008.pdf>.
- Audus KL. Controlling drug delivery across the placenta. *Eur J Pharm Sci*. 1999; 8:161–165. [PubMed: 10379038]
- Banerjee, S.; Carlin, BP.; Gelfand, AE. Monographs on Statistics & Applied Probability. Chapman & Hall/CRC; 2003. Hierarchical modeling and analysis for spatial data..
- Cai B, Lawson AB, McDermott S, Aelion CM. Variable selection for spatial latent predictors under Bayesian spatial model. *Stat Model*. 2011; 11:535–555.
- Cai B, Lawson AB, McDermott S, Aelion CM. Bayesian spatial change-point model for intellectual disability analysis. *Stat Methods Med Res*. 2012 (In press).
- Carroll, R.; Ruppert, D.; Stefanski, LA.; Crainiceanu, C. Measurement error in nonlinear models: a modern perspective. 2nd edn.. Chapman and Hall; New York: 2006.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for mercury. 1999. <http://www.atsdr.cdc.gov/toxprofiles/tp46.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for copper. 2004. <http://www.atsdr.cdc.gov/ToxProfiles/tp132.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for nickel. 2005. <http://www.atsdr.cdc.gov/ToxProfiles/tp15.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for arsenic. 2007a. <http://www.atsdr.cdc.gov/toxprofiles/tp2.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for barium. 2007b. <http://www.atsdr.cdc.gov/ToxProfiles/tp24.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Toxicology profile for chromium. 2012. <http://www.atsdr.cdc.gov/toxprofiles/tp7.pdf>.
- Centers for Disease Control and Prevention. [07 March 2014] Intellectual Disability Fact Sheet. 2014. http://www.cdc.gov/ncbddd/actearly/pdf/parents_pdfs/IntellectualDisability.pdf.
- Collipp PJ, Chen SY, Maitinsky S. Manganese in infant formulas and learning disability. *Ann Nutr Metab*. 1983; 27:488–494. [PubMed: 6651226]
- Cook AJ, Gold DR, Li Y. Spatial cluster detection for repeatedly measured outcomes while accounting for residential history. *Biom J Biom Z*. 2009; 51:801–818.
- Davis HT, Aelion CM, McDermott S, Lawson AB. Identifying natural and anthropogenic sources of metals in urban and rural soils using GIS-based data, PCA, and spatial interpolation. *Environ Pollut*. 2009; 157:2378–2385. [PubMed: 19361902]
- Diggle PJ, Ribeiro PJ Jr. Model-based geostatistics. Springer, New York Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal*. 2007; 1:515–534.
- Goldman LR, Koduru S. Chemicals in the environment and developmental toxicity to children: a public health and policy perspective. *Environ Health Perspect*. 2000; 108(Suppl 3):443–448. [PubMed: 10852843]

- Groom KM, Poppe KK, North RA, McCowan LM. Small-for-gestational-age infants classified by customized or population birthweight centiles: impact of gestational age at delivery. *Am J Obstet Gynecol.* 2007; 197(3):239, e1–5. [PubMed: 17826403]
- Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. Measurement error caused by spatial misalignment in environmental epidemiology. *Biostatistics.* 2009; 10:258–274. [PubMed: 18927119]
- Hoffman K, Webster TF, Weinberg JM, Aschengrau A, Janulewicz PA, White RF, Vieira VM. Spatial analysis of learning and developmental disorders in upper Cape Cod, Massachusetts using generalized additive models. *Int J Health Geogr.* 2010; 9:7. [PubMed: 20152039]
- Kim JI, Lawson AB, McDermott S, Aelion CM. Variable selection for spatial random field predictors under a Bayesian mixed hierarchical spatial model. *Spat Spatio-Temporal Epidemiol.* 2009; 1:95–102.
- Kim JI, Lawson AB, McDermott S, Aelion CM. Bayesian spatial modeling of disease risk in relation to multivariate environmental risk fields. *Stat Med.* 2010; 29:142–157. [PubMed: 19904772]
- Lawson AB. Disease cluster detection: a critique and a Bayesian proposal. *Stat Med.* 2006; 25:897–916. [PubMed: 16453377]
- Lawson, AB. Bayesian disease mapping: hierarchical modeling in spatial epidemiology. Chapman and Hall/CRC; 2013.
- Li X, Lee S, Wong S, Shi W, Thornton I. The study of metal contamination in urban soils of Hong Kong using a GIS-based approach. *Environ Pollut.* 2004; 129:113–124. [PubMed: 14749075]
- Liu Y, McDermott S, Lawson A, Aelion CM. The relationship between mental retardation and developmental delays in children and the levels of arsenic, mercury and lead in soil samples taken near their mother's residence during pregnancy. *Int J Hyg Environ Health.* 2010; 213:116–123. [PubMed: 20045663]
- Manjourides J, Pagano M. Improving the power of chronic disease surveillance by incorporating residential history. *Stat Med.* 2011; 30:2222–2233. [PubMed: 21563208]
- McDermott, S.; Durkin, MS.; Schupf, N.; Stein, ZA. Epidemiology and etiology of mental retardation.. In: Jacobson, JW.; Mulick, JA.; Rojahn, J., editors. *Handbook of intellectual and developmental disabilities.* Springer; New York: 2007.
- McDermott S, Wu J, Cai B, Lawson A, Aelion CM. Probability of intellectual disability is associated with soil concentrations of arsenic and lead. *Chemosphere.* 2011; 84:31–38. [PubMed: 21450328]
- McDermott S, Bao W, Aelion CM, Cai B, Lawson A. When are fetuses and young children most susceptible to soil metal concentrations of arsenic, lead and mercury? *Spat Spatio-Temporal Epidemiol.* 2012; 3:265–272.
- National Dissemination Center for Children with Disabilities. [07 March 2014] 2014. <http://nichcy.org>.
- Paciorek CJ. Combining spatial information sources while accounting for systematic errors in proxies. *J R Stat Soc (Ser C).* 2012; 61:429–451.
- Pless, IB. *The epidemiology of childhood disorders.* Oxford University Press; New York: 1994.
- Spiegelhalter D, Best N, Carlin B, Van der Linde A. Bayesian measures of model complexity and fit. *J Roy Stat Soc Ser B.* 2002; 64:583–616.
- Szpiro AA, Paciorek CJ, Sheppard L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology.* 2011; 22:680–685. [PubMed: 21716114]
- Tong S. Migration bias in ecologic studies. *Eur J Epidemiol.* 2000; 16:365–369. [PubMed: 10959945]
- Trasande L, Schechter CB, Haynes KA, Landrigan PJ. Mental retardation and prenatal methylmercury toxicity. *Am J Ind Med.* 2006; 49:153–158. [PubMed: 16470549]
- U.S. Department of Health And Human Services. [27 March 2012] Toxicological profile for chromium. 2012. <http://www.atsdr.cdc.gov/toxprofiles/tp7.pdf>.
- Vahter M. Effects of arsenic on maternal and fetal health. *Annu Rev Nutr.* 2009; 29:381–399. [PubMed: 19575603]
- Van den Hooven EH, Pierik FH, Van Ratingen SW, Zandveld PY, Meijer EW, Hofman A, Miedema HM, Jaddoe VW, De Kluizenaar Y. Air pollution exposure estimation using dispersion modelling and continuous monitoring data in a prospective birth cohort study in the Netherlands. *Environ Health: Glob Access Sci Sources.* 2012; 11:9.

- Vieira V, Webster T, Aschengrau A, Ozonoff D. A method for spatial analysis of risk in a population-based case-control study. *Int J Hyg Environ Health*. 2002; 205:115–120. [PubMed: 12018004]
- Vieira V, Webster T, Weinberg J, Aschengrau A, Ozonoff D. Spatial analysis of lung, colorectal, and breast cancer on Cape Cod: an application of generalized additive models to case-control data. *Environ Health: Glob Access Sci Sources*. 2005; 4:11.
- Vincenty T. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Surv Rev*. 1975; XXII:88–93.
- Walsh LM, Sumner ME, Keeney DR. Occurrence and distribution of arsenic in soils and plants. *Environ Health Perspect*. 1977; 19:67–71. [PubMed: 908315]
- Wheeler DC, Waller LA, Cozen W, Ward MH. Spatial-temporal analysis of non-Hodgkin lymphoma risk using multiple residential locations. *Spat Spatio-Temporal Epidemiol*. 2012; 3:163–171.
- Zhen H, Lawson AB, McDermott S, Lamichhane AP, Aelion M. A spatial analysis of mental retardation of unknown cause and maternal residence during pregnancy. *Geospat Health*. 2008; 2:173–182. [PubMed: 18686266]

Fig 1.
Mother's movement and location missingness during pregnancy

Table 1

Characteristics of infants, mothers, neighborhood, and soil sample concentrations for mother with and without a child with ID

		ID (n=420)	No ID (n=6,859)	P value
Infant characteristics				
Infant sex	Girl <i>n</i> (%)	136 (3.92 %)	3,337 (96.08 %)	<0.0001 [*]
	Boy <i>n</i> (%)	284 (7.65 %)	3,429 (92.35 %)	
	Missing	0	93	
Weeks of gestation	Mean (sd)	37.55 (3.77)	38.63 (2.48)	<0.0001 [*]
	Missing	21	380	
Maternal characteristics				
Mother's age	Mean (sd)	24.01 (6.25)	22.58 (5.33)	<0.0001 [*]
	Missing	0	93	
Mother's race	White	93 (4.22 %)	2,109 (95.78 %)	0.0005 [*]
	Black	321 (6.59 %)	4,552 (93.41 %)	
	Others	6 (5.56 %)	102 (94.44 %)	
	Missing	0	96	
Alcohol use	Yes	8 (9.41 %)	77 (90.59 %)	0.1700
	No	412 (5.82 %)	6,672 (94.18 %)	
	Missing	0	110	
Tobacco use	Yes	77 (6.53 %)	1,103 (93.47 %)	0.2200
	No	343 (5.62 %)	5,756 (94.38 %)	
	Missing	0	0	
Parity	Mean (sd)	1.15 (1.31)	0.87 (1.13)	<0.0001 [*]
	Missing	0	93	
Neighborhood characteristics				
Mean age of housing years	Mean (sd)	46.46 (11.29)	45.78 (11.70)	0.2500
	Missing	3	14	
Density (Population/sq.) mile)	Mean (sd)	2,903.31 (2,781.88)	2,857.98 (2,690.48)	0.7400
	Missing	2	13	
Chemicals ^{ab} (mg/kg)				
As	Mean (sd)	3.54 (2.37)	3.34 (2.26)	0.0790
Cr	Mean (sd)	19.52 (14.97)	19.80 (14.48)	0.7000
Hg	Mean (sd)	0.16 (0.43)	0.11 (0.35)	0.0086 [*]
Pb	Mean (sd)	81.71 (108.50)	76.31 (119.39)	0.3700
Mn	Mean (sd)	180.91 (184.29)	188.88 (181.29)	0.3800
Ba	Mean (sd)	56.58 (84.37)	51.94 (40.67)	0.0450 [*]
Cu	Mean (sd)	18.048 (14.13)	17.58 (13.69)	0.5000
Ni	Mean (sd)	4.77 (4.77)	4.79 (3.64)	0.9200

^aUse of 10th percentile based on Groom et al. (2007)

^b The metal concentration is the mean value if the mother ever moved

* Indicates statistical significance. *P* values are obtained from a simple logistic regression in a frequentist approach with outcome ID and predictor the variable of interest. The kriged standard deviation was not considered in these models

Table 2

Goodness of fit criteria for retaining metals in the model. Lower DIC values indicate better fit

Model	DIC	pD
As, adjusted ^a	2,933.69	10.84
Cr, adjusted ^a	2,936.66	10.38
Hg, adjusted ^a	2,930.81	10.82
Pb, adjusted ^a	2,936.76	10.88
Mn, adjusted ^a	2,936.32	10.32
Ba, adjusted ^a	2,932.9	11.04
Cu, adjusted ^a	2,935.09	8.65
Ni, adjusted ^a	2,936.43	10.61
As, Hg adjusted ^b	2,923.52	6.52

^a Adjusted for infant sex, alcohol, black race,, smoking, building age of the house, density, parity, mother's age, and weeks

^b Adjusted for infant sex, black race, parity, mother's age, and weeks

Table 3Importance parameters logistic regression results from the adjusted model^a

Chemical ^b	β_0 (95 % CI)	β_1 (95 % CI)	w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	w ₇	w ₈	w ₉
As	-3.69 (-3.99, -3.43) *	0.013 (0.00074, 0.038) *	0.43	0.45	0.44	0.46	0.49	0.47	0.47	0.48	0.47
Hg		0.11 (0.021, 0.39) *	0.49	0.48	0.53	0.52	0.40	0.40	0.37	0.38	0.37

^a Adjusted for infant sex, black race, parity, mother's age, and weeks^b Chemicals measured in mg/kg

* Indicates that the credible interval does not include zero

Table 4

Results from the full-adjusted model. (Tables 3 and 4 include estimates from the same model)

Variable	β_2 (95 % CI)
Infant sex (boy vs. girl)	0.69 (0.48, 0.90) *
Black race (vs. white and others)	0.30 (0.056, 0.56) *
Mother's age **	0.19 (0.092, 0.29) *
Weeks	-0.30 (-0.38, -0.22) *

* Indicates that the credible interval does not include zero

** Mother's age variable has been standardized and therefore β_2 represents the effect per standard deviation

Table 5

Simulation results

	β_0 (95%CI)	β_1 (95 % CI)	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
Model 1	0.047 (-0.11, 0.21)	2.57 (2.54, 2.61) *	0.53	0.53	0.55	0.52	0.53	0.99	0.54	0.54	0.53
Month 6 peak metal exposure											
Model 2	-0.026 (-0.19, 0.13)	0.079 (-2.44, 2.27)	0.50	0.49	0.50	0.49	0.49	0.49	0.50	0.50	0.51
Similar metal exposure											

* Indicates that the credible interval does not include zero