Published in final edited form as:

Epidemiology. 2013 November ; 24(6): . doi:10.1097/EDE.0b013e3182a671e4.

Lipid Adjustment for Chemical Exposures: Accounting for Concomitant Variables

Daniel Li¹, Matthew P. Longnecker², and David B. Dunson¹

¹Department of Statistical Science, Duke University

²Epidemiology Branch, National Institute of Environmental Health Sciences, NIH/DHHS

Abstract

Background—Some environmental chemical exposures are lipophilic and need to be adjusted by serum lipid levels before data analyses. There are currently various strategies that attempt to account for this problem, but all have their drawbacks. To address such concerns, we propose a new method that uses Box-Cox transformations and a simple Bayesian hierarchical model to adjust for lipophilic chemical exposures.

Methods—We compared our Box-Cox method to existing methods. We ran simulation studies in which increasing levels of lipid-adjusted chemical exposure did and did not increase the odds of having a disease, and we looked at both single-exposure and multiple-exposures cases. We also analyzed an epidemiology dataset that examined the effects of various chemical exposures on the risk of birth defects.

Results—Compared with existing methods, our Box-Cox method produced unbiased estimates, good coverage, similar power, and lower type-I error rates. This was the case in both single- and multiple-exposure simulation studies. Results from analysis of the birth-defect data differed from results using existing methods.

Conclusion—Our Box-Cox method is a novel and intuitive way to account for the lipophilic nature of certain chemical exposures. It addresses some of the problems with existing methods, is easily extendable to multiple exposures, and can be used in any analyses that involve concomitant variables.

A concomitant variable is a nonconfounding covariate that, if included in a data analysis, will improve the precision of the estimate of interest.¹ Consider body mass index (BMI), the usual surrogate measure of adiposity. A better measure than weight alone, BMI accounts for how height might influence the effects of weight on a health response – even though height alone may not have any effect.² Other examples of concomitant variables include modifying the dose of a drug in relation to body weight, and in our case, modifying lipophilic chemical concentrations in serum by serum lipid levels.

The analysis of exposure to lipophilic chemicals in relation to potential human health risks poses pressing methodological challenges, as different methods can produce conflicting results. For example, in a recent study of plasma polybrominated diphenyl ether (PBDE) and

Corresponding Author David B. Dunson, Duke University Box 90251, Durham, NC 27708-0251 USA, (919) 260-6615 (P), (919) 684-8594 (F), dunson@stat.duke.edu.

Conflicts of Interest

The authors state no conflict of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (<http://www.epidem.com/>). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the author.

thyroid hormone, free thyroid hormone was associated with higher PBDE levels when PBDE was expressed on a lipid basis, but not when PBDEs were expressed on a wet-weight basis.³ Such differences may be the result of short-term effects such as recent eating.⁴

There has been much debate about how to best deal with lipophilic chemical exposures. One common approach is to express the serum exposure and lipid concentrations as a ratio prior to statistical analysis.⁵ Another approach is to treat serum lipid levels as a covariate in regression analysis.⁶ The first has the theoretical advantage of expressing the serum exposure concentration so it is comparable with that in tissue lipids throughout physiologic compartments in the body. However, a drawback is the imperfect correlation between lipid-based measures in plasma and other issues,^{7,8} which may stem from some lipophilic materials in plasma being bound to albumin rather than being completely sequestered in lipids.⁹ A second drawback is that when an incorrect chemical-to-lipid ratio is chosen, this approach provides biased results under a number of causal models.¹⁰ An advantage of the second approach is that it can accommodate the fact that the correct ratio measure is unknown.¹¹ However, the second approach falls short because it assumes that the exposure effect will be the same regardless of a subject's serum lipid concentrations. Other approaches include using unadjusted wet-weight values, and using the residuals of modeling serum lipid levels by serum exposure levels as a covariate.¹² These approaches are less common and have their own drawbacks.¹⁰

Past studies also have suggested that different methods perform better under different causal structures.¹⁰ For example, if serum lipids are known to directly affect the health response, adding lipids as a covariate would minimize bias in estimating the exposure effect. If serum lipids are known not to affect the health response or exposure, it would be best to not include lipids at all.

We aim to provide a general approach for lipid adjustment under the simplifying assumption that lipids are concomitant variables with no direct causal effect on the health response. When this assumption is substantially violated, it may be more reasonable to adjust for lipids as a covariate in assessing the health effect of an exposure.¹⁰ However, we assume that the direct effect of lipids is typically small, so that lipids act predominately as a concomitant variable. Under this assumption, we propose a new method that uses Box-Cox transformations and a Bayesian hierarchical model, and we compare it to existing methods in simulation studies and a sample analysis.

METHODS

Model Structure

Consider a subject i . Suppose this subject has a conditional probability p_i of having a disease given certain exposure levels and confounding variable measurements. Furthermore, suppose this subject also has a continuous serum chemical exposure concentration x_i . In predicting the conditional odds of having this disease, a natural model choice is a logistic regression:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_i \quad (1)$$

where β_0 is the intercept and β_1 is the exposure slope. We can also include an adjustment for possible confounders $z_i = (z_{i1}, \dots, z_{iq})$ by adding z_i to (1), but we leave off such adjustments here for simplicity.

If the chemical exposure of interest is lipophilic, we should also find some way to incorporate subject i 's total serum lipid concentration, which we will call s_i to adjust the exposure concentration x_i . An intuitive idea is to replace x_i with x_i^* , which we will define as:

$$x_i^* = x_i / g(s_i) \quad (2)$$

where $g(s_i)$ is some unknown function of s_i . Naturally, we can then modify equation (1) and predict the conditional odds of having a disease using:

$$\text{logit}(p_i) = \beta_0^* + \beta_1^* x_i^* \quad (3)$$

where β_0^* is the new intercept and β_1^* is the adjusted exposure's slope.

Use of the ratio approach in equation (2) is biologically motivated, as we want to express the serum level in units equivalent to those in the target tissue. To obtain a more accurate measurement of the actual biological effect of the chemical exposure, we need to allow the chemical dose x_i to be shifted by s_i in some way, $g(s_i)$, to obtain x_i^* . By adjusting for lipid concentrations in this way, x_i^* will ideally be a better predictor to include in the analysis than x_i . This adjustment accounts for the fact that a subject's serum lipid level does not directly contribute to the probability of having a disease, although it can have an indirect effect, as illustrated in Figure 1a. Note that with Figure 1a, we are not implying that serum lipids are always associated with the disease, but simply that serum lipids may sometimes have an effect on what the true adjusted exposure should be. Breaking from conventional directed acyclic graph (DAG) notation, Figure 1b better illustrates this relationship. Furthermore, it is important to realize that this kind of adjustment is already commonly used. For example, in the calculation of BMI,² if height is equal to h , the adjustment in equation (2) is $g(h) = h^2$. As we will also show later, many existing methods employ some form of this adjustment.

Box-Cox Transformation

To estimate the adjustment function, we propose to let $g(s_i)$ correspond to a Box-Cox transformation so that $g(s_i; \lambda) = (s_i^\lambda - 1) / \lambda$ for $\lambda \neq 0$ and $g(s_i; 0) = \log(s_i)$ for $\lambda = 0$. We choose to use a Box-Cox transformation because it is a common but simple class of transformation functions that can model a variety of $g(s_i)$ effects.^{13,14} We are assuming that for a given chemical exposure, there is a single correct but unknown value in which lipids will be properly accounted for. To find the appropriate Box-Cox value to estimate $g(s_i)$, we can calculate and plot the log-likelihoods for various values of λ , as shown in Figure 2. Values of λ that produce larger log-likelihoods can be interpreted as providing better explanations of the data, and the associated Box-Cox transformations can be interpreted as providing better estimations of the function $g(s_i)$.

Allowing λ to vary gives our approach greater flexibility. Consider the scenario in which serum lipids do not affect the health outcome in any way. The standard method of dividing by serum lipid levels mentioned above would be expected to perform poorly.¹⁰ However, our approach will deal with this problem by having the log likelihood peak near values that minimize the effects of the serum lipids. More specifically, these values of λ should produce an x_i^* that is practically identical to x_i , equivalent to not including lipids in the analysis in the first place.

We can further generalize this idea using a hierarchical model. Suppose one is interested in analyzing the effects of multiple exposures all at once, with each exposure having its own value. One problem with the profile likelihood approach above is that, as the number of exposures in the analysis increases linearly, the number of combination grid points

increases exponentially. Another problem is that directly plugging in estimates of β_1 removes from the analysis all uncertainty associated with β_1 . With a hierarchical model, not only can we use the same algorithm to estimate the unknown β_1 values, but we can incorporate the uncertainty from estimating these β_1 values into the confidence intervals. Intuitively, this hierarchical model borrows information from the various exposures and concentrates mass on the β_1 values that allow for the best prediction of the health response.

As an aside, we note that the adjustment x_i^* need not be limited to $x_i/g(s_i)$. A more general and flexible approach would define $x_i^* = f(x_i, s_i)$ where f is a bivariate function. Additional methods may also make $g(s_i)$ more general and not limit it to a Box-Cox transformation. However, with these more general approaches, it is harder to estimate the proper adjustment function, and the results are frequently less interpretable.

Implementation

Within a Bayesian context, our approach is easy to implement. Suppose we start with one of the many Markov Chain Monte Carlo (MCMC) algorithms available for Bayesian logistic regression; such algorithms are routinely implemented in SAS (SAS Institute Inc., Cary, North Carolina), WinBUGS (The BUGS Project, Cambridge, London, England), and R (The R Foundation for Statistical Computing, Vienna, Austria). We can implement our proposed method by adding two sampling steps to an existing MCMC algorithm. These steps will sample from known distributions and be computationally tractable, and because of the Bayesian context, simple probability statements can be attached to inferences. One common criticism of Bayesian methods is the use of subjective prior distributions. To address this, we utilize non-informative priors so that the results of analyses are determined solely by the data. R code from the authors is also available to aid in implementation, and additional technical details can be found in the appendix.

It is worth noting that the interpretation of the exposure effect using our method may appear rather ambiguous at first. Because the units of β_1^* are not known, the interpretation of β_1^* is not as clear as the interpretation of β_1 . However, if we standardize x_i^* so it has mean 0 and standard deviation 1, we can interpret the effect of β_1^* per standard deviation increase in the adjusted predictor. Furthermore, if one is interested in making $x_i^* = \log(x_i/g(s_i))$ to improve modeling, the interpretation of β_1^* would be the same as that of β_1 when using $\log(x_i)$.

Comparison with Existing Methods

We compare our proposed method to three existing methods. The “no adjustment method” involves no lipid adjustment of the chemical exposure. This method is equivalent to letting $g(s_i) = 1$, and can be expected to perform reasonably well if the chemical exposure is not affected by serum lipid concentrations. The second method, the “ratio method”, involves dividing the chemical exposure by some known power of the serum lipid concentration. This is equivalent to letting $g(s_i) = s_i^m$, with m being a predetermined constant. Note that if $m = 0$, this method is the same as the no adjustment method. The third method, the “covariate method”, adds total serum lipids as a covariate in the analysis:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_i + \beta_2 s_i \quad (4)$$

where β_0 is the intercept, β_1 is the exposure variable's slope, β_2 and is the lipid variable's slope.

Single-Exposure Simulation Study

We analyzed data from a reproductive epidemiology study nested in the US Collaborative Perinatal Project.¹⁵ This study collected data on pregnancy outcomes, demographic factors, and levels of exposure to a wide variety of environmental contaminants. Exposure variables included γ -hexachlorocyclohexane (γ -HCH), dieldrin, hexachlorobenzene (HCB), heptachlor epoxide, oxychlorodane, dichlorodiphenyldichloroethylene (DDE), dichlorodiphenyltrichloroethane (DDT), trans-nonachlor, and polychlorinated biphenyls (PCB). The health outcomes were cryptorchidism, hypospadias, or no birth defect.

We performed a simulation study in which the proper lipid adjustment and exposure effects were known. We first considered two single-exposure cases, each with a single chemical exposure and a single confounder. For the first case, we considered a scenario in which the alternative hypothesis was true, in that an increase in the chemical exposure led to an increase in the odds of having the disease outcome. For the second case, we considered a scenario in which the null hypothesis was true, in that there was no association between the chemical exposure and disease outcome. We compared our method with the three competing methods described above. For each method, we generated 1000 data sets and compared biases, Type I errors, power, and mean squared errors.

To generate our simulated data, we used total PCB serum levels and an indicator variable for white race, as found in the sample data. We chose to use total PCB concentration as our exposure because target tissue PCBs depend on serum lipid concentrations.^{5, 16–17} To make exposure adjustments, we approximated total serum lipid levels from total cholesterol and triglyceride measurements using a formula from Phillips et al.¹⁸ Using equation (3), we fixed the coefficients for each predictor and defined the $g(s_j)$ to be a Box-Cox transformation with $\lambda = 1$. Simulation analyses were also run with $g(s_j) = s_j$, but the results were practically identical to those with the Box-Cox transformation, as s_j and $s_j - 1$ are almost the same for large values of s_j . Therefore, only the former results are presented. We also ran the single-exposure simulation with the true λ value equal to 2, but the results were practically the same as those for λ equal to 1 and are omitted as well.

To obtain the adjusted predictor using our method, we first used a Box-Cox transformation to adjust for serum lipids. Next, to allow for direct comparisons between the different methods, we log-transformed the Box-Cox adjusted predictor to obtain our final adjusted predictor. We note that while chemical exposures are frequently log-transformed in analyses, there are no assumptions placed on the distribution of the covariates in running logistic regressions. Log transformation may be considered if they improve model accuracy, but skewed predictors alone do not violate any major assumptions. We proceed with log transformations only because they facilitated consistent interpretation of simulation results across methods.

Multiple-Exposures Simulation Study

We also performed a simulation study in which the odds ratios (OR) for three exposures were known. The purpose of this simulation was to illustrate that our Box-Cox method can be extended to analyzing multiple exposures. The effects of a confounder, two significant exposures, and one non-significant exposure were fixed beforehand. For each method, we ran 1000 data sets and compared biases, Type I errors, power, and mean squared errors.

We used total PCB serum levels, γ -HCH serum levels, HCB serum levels, and an indicator variable for white race to generate the simulated data. Total PCB serum levels, γ -HCH serum levels, and an indicator variable for white race were fixed beforehand to increase the odds of having the disease outcome. HCB was set to not affect the odds. Using equation (3), we fixed the coefficients for each predictor and defined $g(s_j)$ for each exposure to be a Box-

Cox transformation with $m = 1$ or 2 . The adjusted exposure serum levels were also log transformed as described above.

Birth-Defect Data Analysis

We analyzed the ability of each exposure to predict the occurrence of hypospadias and cryptorchidism separately in male births. To maintain consistent interpretations among all exposures and methods, we applied a log transformation to all exposures during analyses. This included setting $x_i^* = \log(x_i / g(s_i))$ when using our proposed method. Additional variables were assessed as potential confounders using the change-in-estimate method,^{19–20} beginning with the inclusion of all variables and removal in a stepwise manner. If the OR of the highest-to-lowest exposure level OR effects changed by $\geq 15\%$, the factor was included in analyses. Potential confounding variables included maternal age, parity, smoking status, maternal pre-pregnancy BMI, socioeconomic index, race, and study center. An indicator variable for white race changed the OR by $\geq 15\%$ for all exposures except total PCBs. An indicator variable for center 55 changed the OR by $\geq 15\%$ for α -HCH, HCB, heptachloro epoxide, and total PCBs. All other variables changed the OR by $<15\%$.

RESULTS

Single-Exposure Simulation Study

Table 1 illustrates the results of the single-exposure simulation study. To calculate the OR, we looked at the effect of a 2-fold increase in chemical exposure. To calculate bias, we took the mean simulated OR for each method and subtracted the true OR. To calculate the mean square error (MSE), we calculated the variance of the simulated ORs and added the square of the bias. For the alternative-hypothesis scenario, our method compares favorably with existing methods. Our approach produced unbiased estimates for the exposure OR effect (OR bias = 0.01, MSE = 0.07), had good coverage, and had the second greatest power (Coverage = 0.95, Power = 0.88). As expected, the ratio method with the “correct” adjustment produced unbiased estimates, had similar coverage, and had the greatest power. However, with the incorrect adjustment, the ratio method produced biased estimates, while coverage and power decreased. To clarify, though the ratio adjustment and Box-Cox adjustment with $m = 1$ are technically different (because s_i takes on large values in this analysis) the two adjustments are practically the same. For the covariate method, estimates are unbiased, and coverage and power are good. Lastly, though the no-adjustment method had good coverage, it produced biased estimates and had lower coverage and power.

Under the null case, our method improved upon existing methods. Though all the existing methods produced unbiased estimates, high coverage, and low type-I error rates (type I error = 0.05), our method had the best coverage and type-I error rates (type I error = 0.01) while also producing unbiased estimates. One can also consider how our method estimated β . Figure 3 shows the posterior distribution for β under the single-exposure alternative case with the true β value equal to 1. (For the remainder of the text, we omit additional posterior distribution plots of β for simplicity in presentation.)

Multiple-Exposures Case

Table 2 illustrates the results of the multiple-exposures simulation study. With PCB exposure, our method performed well compared with existing methods. Estimates are relatively unbiased (OR Bias = -0.02 , MSE = 0.07), and coverage and power are good (Coverage = 0.92, Power = 0.81). The ratio method with the correct adjustment produced unbiased estimates, good coverage, and the best power. However, with an incorrect adjustment, the ratio method produced biased estimates and lower coverage and power. The covariate method produced unbiased estimates, coverage, and power similar to our method.

The no-adjustment method produced biased estimates and smaller coverage and power. Similar results were obtained with the weaker -HCH exposure, with our proposed method producing unbiased estimates (OR bias = -0.01 , RMSE = 0.03), good coverage, and relatively good power (Coverage = 0.95 , Power = 0.19).

The results of the HCB exposure are analogous to those of the single-exposure case under the null hypothesis. In looking at the HCB exposure, our method showed an improvement over the existing methods, producing relatively unbiased estimates (OR bias = -0.04 , MSE = 0.02), the best coverage, and the lowest type-I error rate (type I error = 0.03). The no-adjustment and ratio methods also produced relatively unbiased estimates, and had good coverage and low type-I error rates (type I error = 0.05). Surprisingly, the covariate method performed the worst, as estimates were biased, coverage the lowest, and type-I error rates the highest (type I error = 0.18).

Birth-Defect Data Analysis

Table 3 shows the results of the effects of various chemical exposures on cryptorchidism. The no-adjustment method suggests that none of the investigated chemical exposures are associated with cryptorchidism. The covariate method suggests there is an association between DDT and cryptorchidism, while the ratio method suggests DDT and PCB levels are associated with cryptorchidism. Looking at 95% central credible intervals, our Box-Cox method suggests DDT and total PCB levels are associated with cryptorchidism, just like the ratio method.

Table 4 shows the results of the effects of various chemical exposures on hypospadias. The no-adjustment method led to the conclusion that none of the investigated chemical exposures are associated cryptorchidism, while the covariate method suggests that oxychlorane and PCB levels are associated with hypospadias. The ratio method suggests that heptachloro epoxide, oxychlorane, and PCB and are associated with hypospadias. Looking at 95% central credible intervals, our Box-Cox method suggests oxychlorane and PCB levels are associated with and cryptorchidism, this time like the covariate method.

In analyzing both birth defects, posterior distributions of were found to peak at different values for different chemical exposures. This is to be expected, as different chemicals are expected have different adjustment factors. Aside from this, posterior distributions were similar to those presented in figure 3 and are consequently omitted.

DISCUSSION

We have proposed a new method to adjust for concomitant variables. Our method seeks to estimate the correct adjustment with Box-Cox transformations by using a likelihood maximizing approach, and our method can be extended to analyzing multiple exposures with a hierarchical model. Our method implements an MCMC algorithm that draws from known distributions to provide computational tractability, and is also able to incorporate the uncertainty of estimating the correct adjustment factor in constructing confidence intervals. Our single-exposure simulation studies indicate that our proposed method has good coverage and produces unbiased estimates for exposures in which either the null or alternative are true. For exposures in which the alternative is true, our method's power is comparable to the best of the currently existing methods, and for exposures in which the null is true, our method has the lowest type-I error rate. Similar results were observed in the multiple-exposures simulation study.

Our simulation results also reveal some drawbacks with the existing methods. First, while the ratio method produced the best results when the correct adjustment factor was chosen,

there are some risks with using this approach. The exact lipid adjustment factor is not known for many chemical classes,^{7,21} and even if the exact chemical-class factor were known, the proper adjustment may vary among various health responses. In addition, our simulation results show that when the incorrect lipid adjustment is used, OR estimates are subject to bias, and both coverage and power decrease. Interestingly, our simulation results also show that though the covariate method performed well in the single-exposure simulation studies, it had the highest type-I error rate in the multiple exposures simulation studies. To emphasize, the covariate method performed even worse than the ratio method with the incorrect adjustment. This is a significant drawback because there is an increasing realization of the importance of multi-pollutant approaches in epidemiology and toxicology.^{22–24} Lastly, the simulation study results show that the no-adjustment method produces biased estimates and has less power when adjustments for serum lipids need to be made.

From our sample analyses, we see that different methods can produce different conclusions about the effects of chemical exposures, further emphasizing the importance of properly accounting for serum lipids. These differences in results may be due to the inherent structure of each method. For example, the ratio method fixes the adjustment of serum lipids, while the covariate method assumes serum lipids may have a direct effect on the health response. From the sample analyses, we also see that our method produces interpretations of the exposure effect similar to those of the existing methods. Therefore, not only does our method improve upon the existing methods in some areas, but it also produces results that are just as simple to interpret.

This method can be used in many settings. The ability to implement this method by adding a few additional sampling steps to any existing MCMC algorithm allows our method to be applied in analyses besides logistic regressions. The concomitant variable to be adjusted for also does not have to be restricted to serum lipid levels. For example, a recent study found that adjustment for weight by height in relation to the risk of cancer may vary depending on the cancer site.²⁵ Our proposed method could be used in this scenario to better estimate how to adjust for the concomitant variable height.

Our method is flexible compared with the existing methods, but simple and interpretable relative to potentially more general methods. It addresses some of the drawbacks of existing methods, improves upon their performance in certain situations, and should be useful in adjusting for serum lipid levels and other concomitant variables.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grant R01 ES 017240 from the National Institute of Environmental Health Sciences (NIEHS) of the National Institutes of Health (NIH) and supported in part by intramural funds from NIEHS. The authors would also like to thank various members of the Duke Statistical Science Department and NIEHS for their contributions.

APPENDIX

Box-Cox Hierarchical Model Details

We expect it to be important to borrow information across exposures in estimating $\{ \beta_j \}$, the correct lipid adjustment factor for exposure j . We consider the hierarchical model:

$$\lambda_j \sim P, P = \sum_{h=1}^k \pi_h \delta_{\lambda_h^*}, \pi = (\pi_1, \dots, \pi_k)' \sim \text{Diri} \left(\frac{1}{k}, \dots, \frac{1}{k} \right)$$

where $\{\lambda_h^*\}$ is a pre-specified grid of possible values for λ_j with zero included as one of the grid points, and k equal to the number of pre-specified values for λ_j . This form is motivated by the finite approximation to the Dirichlet process introduced by Ishwaran and Zarepour²⁶ and by computational tractability considerations. This model will tend to assign high probability to a few of the λ_h^* values, concentrating on those atoms that lead to an adjustment that best predicts the health response.

With this model, we can include the covariate adjustment using the above hierarchical prior by modifying existing MCMC algorithms for logistic regression to include a few additional sampling steps. Extending the notation from earlier to the multiple exposures case, let x_{ij}^* equal the serum dose level of exposure j for subject i . This can be accomplished by sampling from the full conditional posterior distributions:

$$\Pr(\lambda_j = \lambda_h^*) \propto \pi_h \times \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \text{ with } p_i \text{ calculated from } x_{ij}^* = x_{ij} / g_j(s_i; \lambda_h^*)$$

$$\pi \sim \text{Diri} \left(\frac{1}{k} + \sum_j 1(\lambda_j = \lambda_1^*), \dots, \frac{1}{k} + \sum_j 1(\lambda_j = \lambda_k^*) \right)$$

Running the Model and MCMC Details

We also used a logistic structure proposed by O'Brien and Dunson²⁷ to further improve computational efficiency. Under the one exposure simulations, we ran our MCMC chain for 5,500 iterations, discarding the first 500 draws. For the multiple exposures simulation, we ran our MCMC chain for 11,000 iterations, and discarded the first 1,000 draws. There was no evidence of lack of convergence based on standard diagnostic tests and examination of trace plots. Effective sample sizes for sampled parameters were also above 1,000.

REFERENCES

1. Harring, J. Concomitant variable. In: Salkind, N., editor. Encyclopedia of research design. Thousand Oaks, CA: SAGE Publications, Inc.; 2010. p. 208-210.
2. Centers for Disease Control and Prevention. <http://www.cdc.gov/healthyweight/assessing/bmi/index.html>.
3. Abdelouahab N, Langlois MF, Lavoie L, Corbin F, Pasquier JC, Takser L. Maternal and cord blood thyroid hormone levels and exposure to Polybrominated Diphenyl Ethers and polychlorinated Biphenyls during early pregnancy. Am J Epidemiology. (In Press).
4. Whitcomb BW, Schisterman EF, Buck GM, Weiner JM, Greizerstein H, Kostyniak P. Relative concentrations of organochlorines in adipose tissue and serum among reproductive age women. Environ Toxicol Pharmacol. 2005; 19:203–213. [PubMed: 21783478]
5. Brown JF Jr, Lawton RW. Polychlorinated biphenyl (PCB) partitioning between adipose tissue and serum. Bull Environ Contam Toxicol. 1984; 33:277–280. [PubMed: 6434008]
6. Moysich KB, Ambrosone CB, Vena JE, Shields PG, Mendola P, Kostyniak P, et al. Environmental organochlorine exposure and postmenopausal breast cancer risk. Cancer Epidemiol Biomarkers Prev. 1998; 7:181–188. [PubMed: 9521429]

7. Mussalo-Rauhamaa. Partitioning and levels of neutral organochlorine compounds in human serum, blood cells, and adipose and liver tissue. *Sci Total Environ.* 1991; 103:159–175. [PubMed: 1909054]
8. Stellman SD, Djordjevic MV, Muscat JE, Gong L, Bernstein D, Citron ML, White A, Kemeny M, Busch E, Nafziger AN. Relative abundance of organochlorine pesticides and polychlorinated biphenyls in adipose tissue and serum of women in Long Island, New York. *Cancer Epidemiol Biomarkers Prev.* 1998; 7:489–496. [PubMed: 9641493]
9. Norén K, Westrand C, Karpe F. Distribution of PCB congeners, DDE, hexachlorobenzene, and methylsulfonyl metabolites of PCB and DDE among various fractions of human blood plasma. *Arch Environ Contam Toxicol.* 1999; 37:408–414. [PubMed: 10473799]
10. Schisterman EF, Whitcomb BW, Louis GMB, Louis TA. Lipid adjustment in the analysis of environmental contaminants and human health risks. *Environ Health Perspect.* 2005; 113(7):853–857. [PubMed: 16002372]
11. Helzlsouer KJ, Aldberg AJ, Huang HY, Hoffman SC, Strickland PT, Brock JW, et al. Serum concentrations of organochlorine compounds and the subsequent development of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 1999; 8:525–532. [PubMed: 10385143]
12. Hunter DJ, Hankinson SE, Laden F, Colditz GA, Manson JE, Willett WC, et al. Plasma organochlorine levels and the risk of breast cancer. *N Engl J Med.* 1997; 337:1253–1258. [PubMed: 9345073]
13. Box GEP, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological).* 1964; 26(2):211–252.
14. Sheather, SJ. Transformations. In: Sheather, SJ., editor. *A Modern Approach to Regression with R.* New York: Springer; 2009. p. 76–102.
15. Klebanoff MA. The collaborative perinatal project: A 50-year retrospective. *Paediatr Perinat Epidemiol.* 2009; 23(1):2–8. [PubMed: 19228308]
16. Eyster JT, Humphrey HE, Kimbrough RD. Partitioning of polybrominated biphenyls (PBBs) in serum, adipose tissue, breast milk, placenta, cord blood, biliary fluid, and feces. *Arch Environ Health.* 1983; 38:47–53. [PubMed: 6299210]
17. Guo YL, Emmet EA, Pellizzari ED, Rohde CA. Influence of serum cholesterol and albumin on partitioning of PCB congeners between human serum and adipose tissue. *Toxicol Appl Pharmacol.* 1987; 87:48–56. [PubMed: 3099428]
18. Phillips D, Pirkle J, Burse V, Bernet J, Henderson L, Needham L. Chlorinated hydrocarbon levels in human serum: effects of fasting and feeding. *Arch Environ Contam Toxicol.* 1989; 18:495–500.
19. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol.* 1993; 138:923–936. [PubMed: 8256780]
20. Trabert B, Longnecker MP, Brock JW, Klebanoff MA, McGlynn KA. Maternal pregnancy levels of trans-nonachlor and oxychlordane and prevalence of cryptorchidism and hypospadias in boys. *Environ Health Perspect.* 2012; 120(3):478–482. [PubMed: 21975279]
21. Calvert GM, Willie KK, Sweeney MH, Fingerhut MA, Halperin WE. Evaluation of serum lipid concentrations among U.S. workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Arch Environ Health.* 1996; 51:100–107. [PubMed: 8638959]
22. Carpenter DO, Arcaro K, Spink DC. Understanding the human health effects of chemical mixtures. *Environ Health Perspect.* 2002; 110(1):25–42. [PubMed: 11834461]
23. De Rosa CT, El-Masri HA, Pohl H, Cibulas W, Mumtaz MM. Implications of chemical mixtures in public health practice. *Journ Toxicol Environ Health.* 2004; 7:339–350.
24. Dominici F, Peng RD, Barr CD, Bell ML. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. *Epidemiology.* 2010; 21(2):187–194. [PubMed: 20160561]
25. Kabat GC, Heo M, Miller AB, Rohan TE. Scaling of weight for height in relation to risk of cancer at different sites in a cohort of Canadian women. *Am J Epidemiol.* 2013; 177:93–101. [PubMed: 23139246]
26. Ishwaran H, Zarepour M. Exact and Approximate Sum Representations for the Dirichlet Process. *The Canadian Journal of Statistics.* 2002; 30:269–283.

27. O'Brien SM, Dunson DB. Bayesian Multivariate Logistic Regression. *Biometrics*. 2004; 60:739–746. [PubMed: 15339297]

FIGURE 1.

a. A directed acyclic graph (DAG) that illustrates how serum lipids, s_i , and the serum exposure level, x_i , affect the adjusted exposure level, x_i^* . The adjusted exposure level consequently affects the probability of a specific health response, p_i , and the observation of the health response, y_i . Observed variables are in boxes and unobserved variables are in circles.

b. The arrow from s_i pointing to the edge between x_i and x_i^* is meant to illustrate how s_i is able to modify the effects of x_i without s_i being a cause of y_i . One can think of this as beam of light going through and being filtered by a lens.

FIGURE 2.

A sample plot of the log-likelihoods of predicting a health response by a Box-Cox transformed predictor as λ varies. Each circle represents a specific log-likelihood evaluated for a specific value of λ . The values on the y-axis are not significant, and this figure is used only to illustrate the profile likelihood approach.

FIGURE 3.

Most of the posterior mass concentrates around Λ equal to 1, the true fixed value, showing that our method does well in estimating Λ .

TABLE 1Single Exposure Simulation Results ($n=1403$)

Method	Coverage of 95% CI ^a	Rejection Rate ^b	Odds Ratio Bias	Odds Ratio MSE
Alternative Hypothesis, True =1, True Odds Ratio = 1.60				
No Adjustment	0.93	0.76	-0.10	0.06
Ratio, m=1	0.95	0.91	0.01	0.06
Ratio, m=2	0.72	0.86	-0.22	0.07
Ratio, m=3	0.19	0.76	-0.37	0.15
Covariate	0.95	0.86	0.02	0.07
Box-Cox	0.95	0.88	0.01	0.07
Null Hypothesis, True =1, True Odds Ratio = 1.00				
No Adjustment	0.95	0.05	0.01	0.02
Ratio, m=1	0.95	0.05	0.01	0.02
Ratio, m=2	0.95	0.05	<0.01	0.01
Ratio, m=3	0.95	0.05	<0.01	0.01
Covariate	0.95	0.05	0.02	0.03
Box-Cox	0.99	0.01	<0.01	0.01

^aCI denotes confidence interval or (central) credible interval.^bRejection rate is power under the alternative hypothesis and type-I error rate under the null hypothesis.

TABLE 2

Multiple Exposures Simulation Results ($n=1403$)

Method	Coverage of 95% CI ^a	Rejection Rate ^b	Odds Ratio Bias	Odds Ratio MSE
Total PCB Variable, True =1, True Odds Ratio = 1.60				
No Adjustment	0.88	0.67	-0.15	0.07
Ratio, m=1	0.95	0.87	0.02	0.07
Ratio, m=2	0.93	0.81	-0.06	0.06
Ratio, m=3	0.91	0.74	-0.13	0.06
Covariate	0.95	0.83	0.03	0.07
Hierarchical Box-Cox	0.92	0.81	-0.02	0.07
-HCH Variable, True =2, True Odds Ratio = 1.19				
No Adjustment	0.89	0.09	-0.10	0.03
Ratio, m=2	0.95	0.25	-0.03	0.02
Ratio, m=3	0.83	0.08	-0.12	0.03
Ratio, m=4	0.73	0.04	-0.16	0.04
Covariate	0.95	0.20	-0.02	0.03
Hierarchical Box-Cox	0.95	0.19	-0.01	0.03
HCB Variable, True =1, True Odds Ratio = 1.00				
No Adjustment	0.95	0.05	0.00	0.03
Ratio, m=1	0.95	0.05	0.01	0.03
Ratio, m=2	0.91	0.09	-0.09	0.03
Ratio, m=3	0.86	0.14	-0.12	0.03
Covariate	0.82	0.18	-0.18	0.07
Hierarchical Box-Cox	0.97	0.03	-0.04	0.02

^a CI denotes confidence interval or (central) credible interval^b Rejection rate is power under the alternative hypothesis and type I error rate under the null hypothesis.

Table 3

Cryptorchidism Results

Exposure	OR – (Posterior Mean) Estimate (95% Confidence/Credible Interval)			
	No Adjustment	Ratio	Covariate	Box-Cox
-Hexachlorocyclohexane	1.09 (0.75, 1.50)	1.23 (0.85, 1.70)	1.16 (0.84, 1.66)	1.21 (0.93, 1.61)
Dieldrin	0.73 (0.33, 1.49)	1.20 (0.64, 2.08)	0.85 (0.40, 1.72)	1.12 (0.86, 1.47)
Hexachlorobenzene	0.97 (0.65, 1.50)	1.15 (0.80, 1.61)	1.00 (0.64, 1.58)	1.11 (0.91, 1.43)
Heptachlor Epoxide	1.08 (0.69, 1.60)	1.28 (0.82, 1.85)	1.16 (0.76, 1.75)	1.18 (0.96, 1.66)
Oxychlorane	1.04 (0.62, 1.73)	1.36 (0.84, 2.12)	1.19 (0.68, 2.07)	1.19 (0.97, 1.62)
DDE	1.12 (0.82, 1.56)	1.25 (0.97, 1.69)	1.20 (0.84, 1.62)	1.23 (0.94, 1.78)
DDT	1.30 (0.92, 1.70)	1.46 (1.00, 1.92)	1.41 (1.02, 1.95)	1.42 (1.06, 2.25)
Trans-nonachlor	1.31 (0.44, 3.52)	1.82 (0.91, 3.46)	1.80 (0.69, 5.20)	1.29 (0.93, 1.91)
PCB	1.32 (0.92, 1.94)	1.54 (1.09, 2.25)	1.46 (0.94, 2.13)	1.34 (1.06, 1.94)

Table 4

Hypospadias Results

Exposure	OR – (Posterior Mean) Estimate (95% Confidence/Credible Interval)			
	No Adjustment	Ratio	Covariate	Box-Cox
-Hexachlorocyclohexane	0.91 (0.61, 1.21)	1.01 (0.76, 1.31)	0.95 (0.62, 1.38)	1.03 (0.74, 1.47)
Dieldrin	0.66 (0.34, 1.37)	1.01 (0.59, 1.73)	0.72 (0.39, 1.55)	1.04 (0.68, 1.33)
Hexachlorobenzene	1.10 (0.79, 1.62)	1.20 (0.84, 1.78)	1.13 (0.76, 1.73)	1.13 (0.90, 1.55)
Heptachlor Epoxide	1.37 (0.96, 2.04)	1.50 (1.02, 2.26)	1.47 (0.94, 2.29)	1.31 (0.97, 1.98)
Oxychlorane	1.67 (0.95, 2.88)	1.86 (1.14, 3.04)	1.92 (1.12, 3.35)	1.50 (1.03, 2.59)
DDE	0.93 (0.63, 1.23)	1.00 (0.79, 1.32)	0.97 (0.78, 1.32)	1.05 (0.78, 1.34)
DDT	0.89 (0.61, 1.28)	0.95 (0.77, 1.28)	0.92 (0.65, 1.25)	1.01 (0.66, 1.37)
Trans-nonachlor	1.42 (0.58, 3.84)	1.62 (0.80, 3.04)	1.81 (0.63, 5.29)	1.23 (0.99, 1.92)
PCB	1.47 (0.97, 2.11)	1.62 (1.17, 2.31)	1.59 (1.07, 2.31)	1.41 (1.02, 2.19)