

USING THE SEYCHELLES CHILD DEVELOPMENT STUDY TO CLUSTER MULTIPLE OUTCOMES INTO DOMAINS TO IMPROVE ESTIMATION OF THE OVERALL EFFECT OF MERCURY ON NEURODEVELOPMENT

AMY LALONDE AND TANZY LOVE

Abstract. Environmental exposure effects on human development can be small and difficult to detect due to the nature of observational data. In the Seychelles Child Development Study, researchers examined the effect of prenatal methylmercury exposure using a battery of tests measuring aspects of child development [18, 20]. We build a multiple outcomes model similar to that of the previous analyses (see [18, 20]); however, our multiple outcomes model makes no assumptions of relationships between the testing outcomes. Instead, the nesting of outcomes into domains is a clustering problem we address with a Dirichlet process mixture model implemented through a Bayesian MCMC approach [12]. This model provides inference for the methylmercury exposure effect as well as greater insight into the similarities and differences across the outcomes.

1. INTRODUCTION

Analyzing data on multiple outcomes for a single mechanism of interest arises across various disciplines. Within biomedical research alone, clinical trials utilize multiple endpoints to assess treatment efficacy [13], developmental toxicity studies on laboratory animals to determine exposure effects can include the analysis of multiple ordinal, continuous and binary responses [19], the development of accurate diagnostic measures of birth defects required the consideration of multiple quantitative and subjective measures [2], and epidemiological studies on exposure effects necessitated the analysis of multiple outcome measures [18].

The statistical analysis of such data is often complex, with various methods proposed in the literature exhibiting utility in the aforementioned areas. The more traditional techniques that have been used are Hotelling's T^2 statistic, multiple testing adjustments, two-stage factor analysis, generalized estimating equations [9], and the multivariate analysis of variance. However, these methods lack the ability to comprehensively describe the data in an efficient manner [15]. As such, it

MSC (2010): primary 62F15, 62P12, 62J05.

Keywords: Dirichlet process prior, Bayesian, model-based clustering, multivariate linear models.

The authors wish to thank Sally W. Thurston for many helpful discussions of multiple outcome models and Phil Davidson, Gary Myers, Edwin van Wijngaarden for their helpful comments. This publication was made possible by grant numbers 5-R01-ES08442 and P30 ES 01247 from the National Institute of Environmental Health Sciences (NIEHS).

has often been the aim of the multiple outcomes models to test for global or overall covariate effects. The use of a single model affords efficiency gains and broader insight into outcomes through the use of outcome-specific effects, while also flexibly capitalizing on similarities across subsets of outcomes by introducing a grouping scheme within the overall model. Identifying subgroups of similar outcomes, which are thought of as manifestations of underlying domains or outcome classes, has typically been made a priori. The construction of domains facilitates the borrowing of information across similar measures. As such, our methodology utilizes the information within the data to inform the nesting of outcomes.

When analyzing environmental exposure effects, which are often small and difficult to detect given the observational nature of the data, multiple outcomes models with relatively few assumptions offer the opportunity for improved inference. Our motivation is the Seychelles Child Development Study (SCDS), a longitudinal study aimed at exploring the effect of prenatal methylmercury (MeHg) exposure from maternal fish consumption on the functioning of the central nervous system. Specifically, we analyze 20 testing outcomes evaluated on the cohort of 9-year-old children in the SCDS by fitting a linear mixed model (LMM) that relaxes assumptions made in previous analyses (see [11, 18, 20]). Methylmercury is recognized as a neurotoxin having destructive effects on brain and nervous tissue. Previous analyses explored the effect of MeHg based on the testing outcomes, which were considered as manifestations of four domains of child development: cognitive, memory, motor and behavioral functioning [18, 20]. However, we posit that the complexity of the outcomes may offer the opportunity for model-based clustering techniques that do not place any restrictions on the size and number of clusters to effectively identify and group outcomes into domains. We further believe that such rearrangements will improve model fit diagnostics and estimates of the effects of interest. This work is outlined as follows: Section 2 describes the SCDS, Section 3 details our methodology, and Section 4 presents the results, concluding with a discussion in Section 5.

2. SEYCHELLES CHILDHOOD DEVELOPMENT STUDY

The Seychelles Childhood Development Study, which began enrollment in 1989, is an international research collaboration between the Ministries of Health and Education in Seychelles, the University of Ulster in Northern Ireland, and the University of Rochester in New York, USA. Motivated by previous findings that prenatal exposure to MeHg within seed grain in Iraq harmfully affected child development (see [3]), the SCDS aimed at investigating whether the MeHg within fish had similar adverse effects on a child's neurodevelopment. Fish consumption is the main source of human exposure to MeHg; however, fish constitutes a healthy source of protein rich in nutrients known to be beneficial to brain development despite MeHg's recognition as a neurotoxin [5]. The Seychelles is a group of small islands off the east coast of Africa where the population's fish consumption greatly exceeds that of the United States and many other countries. Widespread accessibility to quality health care and education in the Seychelles reduces the confounding that could occur due to these factors. Hence, the Seychelles offers an ideal center for the observational study.

Beginning in 1989, the first cohort of pregnant mothers was enrolled with the intention of studying their unborn childrens' neurodevelopment. Hair samples were used to measure levels of MeHg in the childbearing women as a surrogate for the prenatal exposure. This cohort of children was evaluated at multiple time-points; however, we focus on the testing at nine years of age. With the aim of fully encompassing a child's neurodevelopment, the battery of tests consisted of multiple widely recognized tests in the neuropsychological literature that are aimed at comparing individuals to the norms dictated by the population as a whole [16]. Myers et al. first analyzed the effect of prenatal MeHg exposure on the outcomes of the tests using separate multivariate regression models, controlling for relevant covariates [11]. Out of the 21 outcomes, only two demonstrated significant associations with MeHg; however, one of the two showed a beneficial effect and the other an adverse effect. Thurston et al. re-analyzed 20 of the outcomes within a single multiple outcomes model to allow for estimation of an overall MeHg effect by nesting outcomes into four broad domains: cognition, memory, motor and behavior [18]. We analyze these same 20 outcomes, summarized in Table 1; however, we use Bayesian clustering techniques to allow the data to inform the optimal nesting of outcomes into a potentially new set of domains. Based on the original multiple outcomes model in [18], we log-transformed six outcomes and subsequently centered and scaled all outcomes to better satisfy the regression assumptions. Seven outcomes were negated to allow positive values to indicate better performance. Furthermore, there were 533 subjects with complete covariate data, and of those, for the 142 subjects missing one or more of the outcome variables, we perform multiple imputation using standard data augmentation methods [17].

Table 1. The table above outlines the 20 outcomes we analyze as nested in [18].

Domain	Outcomes
Cognitive	y_1 -WISC Verbal IQ, y_2 -WISC Performance IQ, y_3 -Boston-Naming Test
Memory	y_4 -CVLT: Recognition, y_5 -CVLT: Immediate Recall, y_6 -CVLT: Short Delay Recall, y_7 -CVLT: Long Delay Recall, y_8 -WISC-R: Digit Span, y_9 -WRAML: Design Memory
Motor	y_{10} -Grooved Pegboard: Dominant Hand, y_{11} -Grooved Pegboard: Non-dominant Hand, y_{12} -Finger tapping: Dominant Hand, y_{13} -Finger tapping: Non-dominant Hand, y_{14} -Trailmaking A, y_{15} -Trailmaking B, y_{16} -Bruininks-Oseretsky, y_{17} -Beery-Buktenica
Behavior	y_{18} -CBCL: Internalizing T Score, y_{19} -CBCL: Externalizing T Score, y_{20} -CTRS Hyperactivity Index

In addition to the MeHg exposure, we control for six covariates, selected based on significance in at least two of the separate regressions on each of the outcomes. These covariates are child's sex and testing age, maternal age and K-BIT (a measure of IQ), HOME score (a measure of stimulation in the home environment), and Hollingshead socioeconomic status, summarized in Table 1. We treat these covariates as domain-specific effects because it is likely the effects are not uniform across domains. In contrast, MeHg is treated as an overall effect with domain- and outcome-specific deviations to provide a means to estimate an overall MeHg effect,

while also allowing for the borrowing of strength across domains and outcomes as all tests aim at measuring brain functioning in some capacity.

Table 2. The table above provides descriptive statistics of the covariates we control for in our multiple outcomes model. These are the same covariates controlled for in Thurston et al.[18] and Xiao et al.[20].

Covariate	Mean (SD)	Range
Child's Age	9.04 (0.31)	8.45 - 10.09
Mother's Age	26.09 (5.78)	14.33 - 44.79
HOME Score	33.56 (5.34)	13 - 47
K-BIT	80.62 (15.09)	45 - 125
Hollingshead Socioeconomic Status	25.55 (11.02)	3 - 58
Male Child	268 (50.3%)	

3. METHODS

Let $\mathbf{y}_{j,d_j} = (y_{j,d_j,1}, \dots, y_{j,d_j,n})^T$ be the vector of centered and scaled observations for all n subjects on the j^{th} testing outcome, assigned to the d_j^{th} domain. The same model formulation applies for all outcomes, $j = 1, \dots, J$ within domains, $d_j \in \{1, \dots, D\}$. Thurston et al. [18] proposed the following model:

$$\mathbf{y}_{j,d_j} = \mathbf{S}_{\mathcal{F}}^T \beta_{\mathcal{F}} + \mathbf{S}_{\mathcal{D}\mathcal{F}} \beta_{\mathcal{D}\mathcal{F},d_j} + \mathbf{S}_{\mathcal{D}} b_{\mathcal{D},d_j} + \mathbf{S}_{\mathcal{O}} b_{\mathcal{O},j} + \mathbf{r} + \mathbf{r}_{\mathcal{D},d_j} + \epsilon_j \quad (3.1)$$

where the matrices, $\mathbf{S}_{\mathcal{F}}$, $\mathbf{S}_{\mathcal{D}\mathcal{F}}$, $\mathbf{S}_{\mathcal{D}}$, and $\mathbf{S}_{\mathcal{O}}$, represent the model covariates included in the overall and domain-specific fixed effects, and in the domain- and outcome-specific random effects, respectively. Furthermore, \mathbf{r} , $\mathbf{r}_{\mathcal{D},d_j}$, and ϵ_j denote the overall and domain-specific random subject effects, and the residual error, respectively. Interest lies in both the overall exposure effect, $\beta_{\mathcal{F}}$, and the total outcome-specific exposure effects, $(\beta_{\mathcal{F}} + b_{\mathcal{D},d_j} + b_{\mathcal{O},j})$, which is the slope of the MeHg exposure variable for each outcome j in domain d_j .

In Thurston et al., the d_j represented an a priori nesting of the j^{th} outcome into the d_j^{th} domain for $d_j \in \{1, \dots, D\}$ where $D = 4$ is fixed at the total number of domains, defined by the cognition, memory, motor and behavioral functions the tests are intended to measure. We form our model identical to Eq. (3.1) except that the d_j are now treated as random. To that end, we incorporate a Dirichlet process mixture model on the d_j , which fits seamlessly within a Bayesian model framework.

Dirichlet process mixture models (DPMM) were first seen in Antoniak [1] and Ferguson [7], with the developments by Escobar and West [6], MacEachern and Müller [10], and others making these practical approaches to data analysis. Neal presented several sampling methods, some of which we utilize here, that further expanded the ease of implementing these models in a standard MCMC [12]. The DPMM is a common modeling choice when dealing with mixture models with an unknown number of components. The mixture is defined as an infinite mixture without any restrictions on the number of components. It has been shown, however, that the total number will grow rather slowly at a rate of $\log(n)$ where n is the total number of items to be clustered, ensuring reasonably sparse results.

To define our DPMM, let $\mathbf{d} = (d_1, \dots, d_J)$ be the vector of domain assignments for the J outcomes, where $d_j = d$ corresponds to the j^{th} outcome's membership in domain d , and let $\boldsymbol{\vartheta} = (\boldsymbol{\theta}, \boldsymbol{\phi})$ where $\boldsymbol{\theta} = (\beta_{\mathcal{F}}, b_{\mathcal{O}}, \mathbf{r}, \mathbf{r}_{\mathcal{D}}, \sigma_{b_{\mathcal{D}}}^2, \sigma_{b_{\mathcal{O}}}^2, \sigma_r^2, \sigma_{\epsilon,1}^2, \dots, \sigma_{\epsilon,J}^2)$ are the non-domain specific parameters. Then, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_D)$ where $\phi_d = (\boldsymbol{\beta}_{\mathcal{D}\mathcal{F},d}, b_{\mathcal{D},d}, \mathbf{r}_{\mathcal{D},d}, \sigma_{r_{\mathcal{D},d}}^2)$ are the parameters defining domain d . We formulate a DPMM on the outcomes as:

$$\begin{aligned} \mathbf{y}_j | \phi_d &\sim F(\phi_d) \\ \phi_d | F_{\phi} &\sim F_{\phi} \\ F_{\phi} &\sim DP(F_{\phi_0}, \alpha) \end{aligned}$$

where $F(\phi_d)$ is the likelihood of the outcomes conditional on their domain assignment, d , and F_{ϕ_0} is the base distribution for the Dirichlet process. In our case, we assume $F(\phi_d)$ is a normal distribution with mean equal to Eq. (3.1) with $d_j = d$ and variance $\sigma_{\epsilon,j}^2 \mathbf{I}_{n \times n}$, and that F_{ϕ_0} has prior density given by

$$\begin{aligned} f_{\phi}(\phi_d) &= f_{\phi}(\boldsymbol{\beta}_{\mathcal{F},\mathcal{D},d}, b_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \sigma_{r_{\mathcal{D},d}}^2) \\ &= f(\boldsymbol{\beta}_{\mathcal{F},\mathcal{D},d}) f(b_{\mathcal{D},d} | \sigma_{b_{\mathcal{D}}}^2) f(\mathbf{r}_{\mathcal{D},d} | \sigma_{r_{\mathcal{D},d}}^2) f(\sigma_{r_{\mathcal{D},d}}^2) \\ &= N_{p_{\mathcal{D}\mathcal{F}}}(\boldsymbol{\beta}_{\mathcal{D}\mathcal{F}0}, \boldsymbol{\Sigma}_{\mathcal{D}\mathcal{F}0}) N_{p_{\mathcal{D}}}(\mathbf{0}, \sigma_{b_{\mathcal{D}}}^2) N_n(\mathbf{0}, \sigma_{r_{\mathcal{D},d}}^2) IG(A_{r_{\mathcal{D}}}, B_{r_{\mathcal{D}}}). \end{aligned}$$

Hyperparameters values for $\boldsymbol{\beta}_{\mathcal{D}\mathcal{F}0}$, $\boldsymbol{\Sigma}_{\mathcal{D}\mathcal{F}0}$, $A_{r_{\mathcal{D}}}$, and $B_{r_{\mathcal{D}}}$ are discussed in the following section on priors.

3.1. Considerations in using a Dirichlet process mixture model

Outcomes are assigned to pre-existing domains or a new domain according to the conditional probabilities computed by

$$\begin{aligned} P(d_j = d \text{ for some } d = d_i \text{ where } i \neq j | d_{-j}, \mathbf{y}_j, \boldsymbol{\phi}) &= b \frac{n_{-j,d}}{J-1+\alpha} F(\mathbf{y}_j, \phi_d) \\ P(d_j \neq d_i \text{ for all } i \neq j | d_{-j}, \mathbf{y}_j, \boldsymbol{\phi}) &= b \frac{\alpha}{J-1+\alpha} \int F(\mathbf{y}_j, \phi) dF_{\phi_0}(\phi) \end{aligned} \quad (3.2)$$

where $n_{-j,d}$ is the number of outcomes in domain d when removing outcome j from consideration, and b is the appropriate normalizing constant to make the probabilities sum to one. While the denominator $J-1+\alpha$ will remain constant and can be absorbed into b , it is useful to see the weighting scheme that applies to each posterior probability. For example in the case of $\alpha = 1$, the ‘‘weights’’ reduce to the proportion of outcomes in each domain when outcome j is placed in its own domain (i.e., $n_{-j,d}/J$ for the pre-existing domains, and $1/J$ for the supposed singleton consisting of only the j^{th} outcome). The likelihood component, $F(\mathbf{y}_j, \phi_d)$, represents a measure of closeness for the j^{th} outcome to the d^{th} domain based on the domain's current parameter values, ϕ_d , which depend largely on its current composition of outcomes. Larger likelihood values indicate greater similarity to the outcomes currently associated with that particular domain. In cases where outcome j does not fit well with any of the current outcome arrangements, the likelihood values will be small for all d , and there may be greater probability of forming a new domain. When considering a new domain, we must formulate

a similar measure of closeness to a supposed new set of parameter values. Integrating the posterior over all possible parameter values, as seen in $\int F(\mathbf{y}_j, \phi) dF_{\phi_0}(\phi)$, delivers this measure of closeness.

The probabilities in Eq. (3.2) can be computed explicitly in the case of conditional conjugacy. In our model, this requires a simplifying assumption that $\sigma_{r_{\mathcal{D},d}}^2$ is constant across all the domains in order to explicitly compute $\int F(\mathbf{y}_j, \phi) dF_{\phi_0}(\phi)$. This is a fairly strong assumption; hence, doing so led to poor model performance. Therefore, we chose to keep the nonconjugate model formulation. Sampling for such cases requires special consideration. Neal presented the “no gaps” algorithm of MacEachern and Müller, which uses a Monte Carlo approximation of the integral [12] to provide straightforward sampling.

Using the probabilities in Eq. (3.2) to directly update the domain assignments in either conjugate or nonconjugate cases leads to incremental updates. Thus, we update one domain assignment at a time as we traverse the sequence of outcomes; however, each potential reassignment will leave the ϕ_d unaltered, yet the “weights” will reflect the rearrangement. This can lead to poor model performance as outcomes can become “stuck” in regions of low posterior probability. To circumvent this issue, we implemented a split-merge procedure as proposed in Jain and Neal, which allowed groups of outcomes to be rearranged in one move with appropriately proposed ϕ_d to accompany a potentially updated \mathbf{d} . These split-merge proposals were accepted based on a Metropolis Hasting probability [8].

3.2. Model Assumptions and Priors

Completing the Bayesian specification of our model, we assume that the random effects, $b_{\mathcal{D},d}$, $b_{\mathcal{O},j}$, $r_{\mathcal{D},d,i}$, are independent of one another and of $\beta_{\mathcal{F}}$, $\beta_{\mathcal{DF},d}$, ϵ_j for all domains, d , and all outcomes, j . Furthermore, we assume $b_{\mathcal{D},d} \sim N(0, \sigma_{b_{\mathcal{D}}}^2)$, $b_{\mathcal{O},j} \sim N(0, \sigma_{b_{\mathcal{O}}}^2)$, $r_{\mathcal{D},d,i} \sim N(0, \sigma_{r_{\mathcal{D},d}}^2)$ and $\epsilon_i \sim N(0, \Sigma_{\epsilon})$, where $\Sigma_{\epsilon} = \text{diag}(\sigma_{\epsilon,1}^2, \dots, \sigma_{\epsilon,J}^2)$ for all $j = 1, \dots, J$, and $d = 1, \dots, D$.

For parameters which are not domain-specific, we select conditionally conjugate priors, setting $\beta_{\mathcal{F}} \sim N(\beta_0, \sigma_0^2)$, $\sigma_{b_{\mathcal{D}}}^2 \sim IG(A_{b_{\mathcal{D}}}, B_{b_{\mathcal{D}}})$, $\sigma_{b_{\mathcal{O}}}^2 \sim IG(A_{b_{\mathcal{O}}}, B_{b_{\mathcal{O}}})$, $\sigma_r^2 \sim IG(A_r, B_r)$ and $\sigma_{\epsilon,1}^2, \dots, \sigma_{\epsilon,J}^2 \sim IG(A_{\epsilon}, B_{\epsilon})$. For the domain-specific parameters, we set $\beta_{\mathcal{DF},d} \sim N(\beta_{\mathcal{DF}0}, \Sigma_{\mathcal{DF}0})$ and $\sigma_{r_{\mathcal{D},d}}^2 \sim IG(A_{r_{\mathcal{D}}}, B_{r_{\mathcal{D}}})$, which appear conditionally conjugate but are not in the case of the DP prior (as previously mentioned). In matching several of our priors to [18], we also set the hyperparameter values to $\beta_0 = 0$, $\sigma_0^{-2} = 0.00001$, $\beta_{\mathcal{DF}0} = \mathbf{0}$, $\Sigma_{\mathcal{DF}0}^{-1} = \text{diag}(0.00001)$, $A_{b_{\mathcal{D}}} = A_{b_{\mathcal{O}}} = A_r = A_{r_{\mathcal{D}}} = A_{\epsilon} = 0.5$, $B_{b_{\mathcal{D}}} = B_{b_{\mathcal{O}}} = B_r = B_{r_{\mathcal{D}}} = 0.00005$ and $B_{\epsilon} = 0.0005$. These values represent weakly informative prior specifications. Our model requires further specification of the DP prior hyperparameter, α , which we set equal to 1.

4. RESULTS

We initiated four MCMC chains from different initial clusterings for our outcomes: all outcomes in separate domains, outcomes assigned to the original a priori nesting in [18], and outcomes assigned sequentially to 8 and then 12 domains. MCMC samplers run for 10,000 iterations, from each of these starting points generated

identical posterior inference for the final nesting of outcomes into domains. The final cluster assignments were determined based on the least squares clustering method proposed in Dahl [4]. Inference for the remaining parameters of interest was generated from an MCMC sampler using Eq. (3.1) with d_j fixed at the posterior nesting choice. A sample of 1,000 draws was generated from a chain run for 200,000 iterations, thinned to every 80 and excluding a burnin of 1,500. Raftery–Lewis convergence diagnostics [14] indicated sufficient draws for estimating the 0.025th quantile for all parameters of interest within ± 0.0125 with 95% probability.

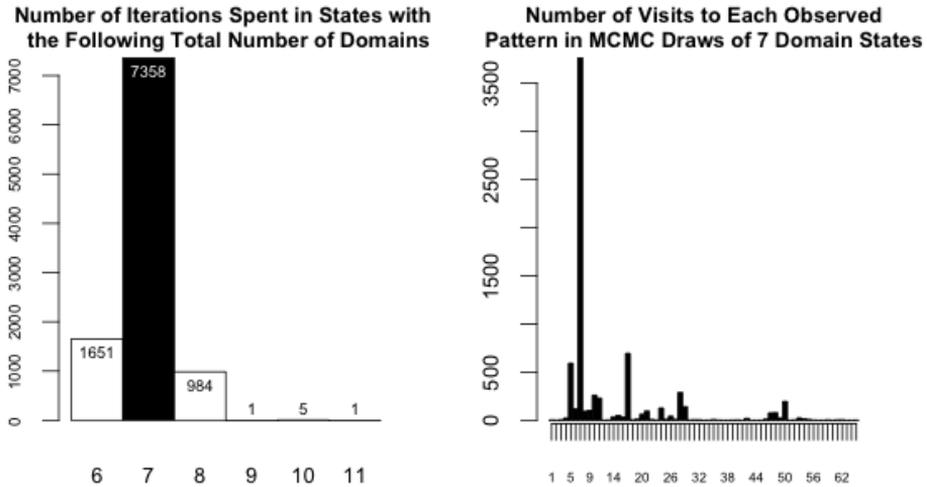


Figure 1. The plot on the left summarizes the posterior distribution of the total number of domains, while the plot on the right summarizes the frequency of MCMC sampler visits to 7 domain states. As expected, the least squares clustering selects the pattern observed most among states with 7 domains.

Figure 1 summarizes the performance of the infinite mixture model in the MCMC sampler used to determine the clustering. There is strong evidence for 7 domains in the plot on the left, while the plot on the right displays strong support for the final clustering selected using the least squares method (the arrangement was visited over 35% of the time in our sampler). Table 3 reveals that the resulting 7-domain posterior clustering retains a grouping similar to the a priori nesting in that only domains 3 and 4 contain outcomes from more than one of the a priori domains. Domain 1 contains only the previously defined cognition outcomes, suggesting agreement with the a priori nesting, and several other domains contain only those outcomes from the same battery of tests (e.g. CVLT test, paired motor tests, and the CBCL tests). Moreover, domains 3 and 4, the only two domains containing outcomes from more than one battery of tests (besides domain 1 which was grouped that way a priori), were also the most unstable domains in the posterior nesting in the sense that their outcomes saw the most movement in the clustering procedure. The estimates of other model parameters suggest that these domains are likely formed due to similarities in the outcome-specific MeHg effect

(see Figure 2) and in the fixed covariate effects. The fixed covariate effects proved to be stronger than the MeHg exposure, demonstrating significance across several domains; however, we do not depict them here as our interest was the overall effect of prenatal MeHg exposure.

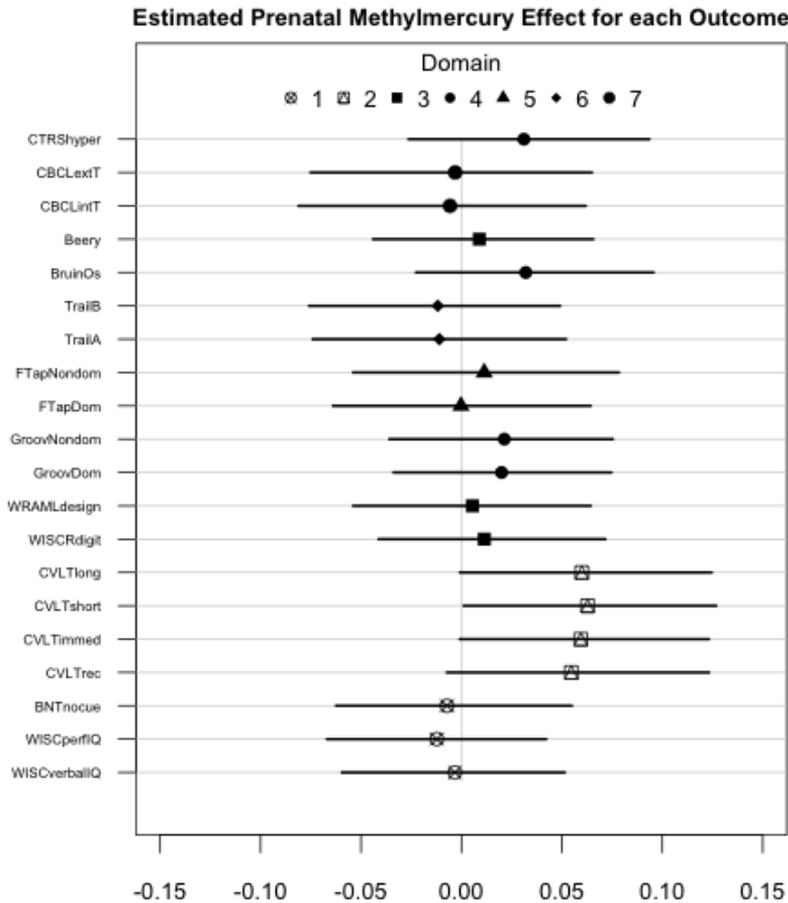


Figure 2. The plot depicts our estimated overall outcome-specific MeHg effects and 95% posterior intervals. We see that most estimated effects show slightly beneficial, albeit nonsignificant, effects of prenatal MeHg exposure.

5. DISCUSSION

We extended the multiple outcomes model of [18], finding a middle ground between the fixed nesting and the partial nesting proposed in [20] with application to the same SCDS data. Our Bayesian model framework incorporates a clustering procedure aimed at nesting outcomes into domains based on similarities in the MeHg exposure effect, the six fixed covariate effects, and random subject-specific effects. The domain-specific parameters provided the information necessary to

Table 3. Posterior nesting of outcomes into domains following the least squares clustering procedure applied to the posterior draws of the domain assignments. Horizontal lines separate the 8 posterior domains, with the second column denoting the original a priori nesting from [18].

Posterior Nesting	A Priori Domain	Outcome	MeHg Effect Estimate	SE	95% Interval
1	Cognition	WISC verbal IQ	-0.003	0.028	(-0.06, 0.05)
1	Cognition	WISC performance IQ	-0.007	0.029	(-0.07, 0.05)
1	Cognition	Boston Naming Test: no cues	-0.003	0.029	(-0.06, 0.06)
2	Memory	CVLT recognition	0.06	0.033	(-0.01, 0.12)
2	Memory	CVLT immediate recall	0.063	0.033	(-0.001, 0.12)
2	Memory	CVLT short delay recall	0.065	0.033	(0.001, 0.13)
2	Memory	CVLT long delay recall	0.065	0.033	(-0.001, 0.13)
3	Memory	WISC-R digit span	0.014	0.028	(-0.04, 0.07)
3	Memory	WRAML design memory	0.011	0.028	(-0.05, 0.06)
3	Motor	Beery-Buktenica	0.013	0.028	(-0.04, 0.07)
4	Motor	Grooved pegboard: dominant	-0.011	0.033	(-0.03, 0.08)
4	Motor	Grooved pegboard: nondominant	-0.01	0.033	(-0.04, 0.08)
4	Motor	Bruininks-Oseretsky	0.079	0.035	(0.02, 0.10)
4	Behavior	CTRS hyperactivity index	0.079	0.036	(-0.03, 0.09)
5	Motor	Finger tapping dominant	0.002	0.034	(-0.06, 0.06)
5	Motor	Finger tapping nondominant	0.012	0.034	(-0.05, 0.08)
6	Motor	Trailmaking A	-0.008	0.032	(-0.07, 0.05)
6	Motor	Trailmaking B	-0.008	0.032	(-0.08, 0.05)
7	Behavior	CBCL: internalizing T score	-0.005	0.037	(-0.08, 0.06)
7	Behavior	CBCL: externalizing T score	-0.003	0.037	(-0.08, 0.07)

group the 20 testing outcomes into seven domains, as opposed to the original four domains. Not surprisingly, several of the domains were comprised solely of outcomes from the same battery of tests. Domains containing outcomes for multiple batteries of tests saw the most variability. Despite the instability of two such domains, our model yields noteworthy improvements in model fit diagnostics.

The nesting of outcomes into domains within multiple outcomes models is intended to allow further borrowing of strength across outcomes within a domain. With fixed a priori nesting of outcomes, we introduce the possibility of missing important relationships in the data. By allowing the data to inform the clustering of outcomes into domains, as we did here, we further improve our estimates of the parameters of interest while also learning more about the relationships between the outcomes.

REFERENCES

- [1] C. E. Antoniak, *Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems*, *Annals of Statistics* **82** (1974), 1152–1174.
- [2] B. A. Coull, J. P. Hobert, L. M. Ryan and L. B. Holmes, *Crossed random effect models for multiple outcomes in a study of teratogenesis*, *Journal of the American Statistical Association* **96** (2001), 1194–1204.
- [3] C. Cox, T. W. Clarkson, D. O. Marsh, L. Amin-Zaki, S. Tikriti and G. G. Myers, *Dose-response analysis of infants prenatally exposed to methyl mercury: An application of a single compartment model to single-strand hair analysis*, *Environmental Research* **49** (1989), 318–332.
- [4] D. B. Dahl, *Model-based clustering for expression data via a Dirichlet process mixture model*, in: K.-A. Do, P. Müller, M. Vannucci (eds.), *Bayesian Inference for Gene Expression and Proteomics*, 2006, 201–218.
- [5] J. L. Daniels, M. P. Longnecker, A. S. Rowland and J. Golding (ALSPAC Study Team, University of Bristol Institute of Child Health), *Fish intake during pregnancy and early cognitive development of offspring*, *Epidemiology* **15** (2004), 394–402.
- [6] M. D. Escobar and M. West, *Bayesian density estimation and inference using mixtures*, *Journal of the American Statistical Association* **90** (1995), 577–588.

- [7] T. S. Ferguson, *Bayesian density estimation by mixtures of normal distributions*, Recent Advances in Statistics, Papers in Honor of Herman Chernoff on his Sixtieth Birthday (1983), 287–302.
- [8] S. Jain, R. M. Neal, *Splitting and merging components of a nonconjugate Dirichlet process mixture model*, Bayesian Analysis **2** (2007), 445–472.
- [9] K. Y. Liang and S. L. Zeger, *Longitudinal data analysis using generalized linear models*, Biometrika **73** (1986), 13–22.
- [10] S. N. MacEachern and P. Müller, *Estimating mixture of Dirichlet process models*, Journal of Computational and Graphical Statistics **7** (1998), 223–238.
- [11] G. J. Myers, P. W. Davidson, C. Cox, C. F. Shamlaye, D. Palumbo, E. Cernichiari, J. Sloane-Reeves, G. Wilding, J. Kost, L. S. Huang and T. W. Clarkson, *Prenatal methylmercury exposure from ocean fish consumption in the Seychelles child development study*, Lancet **361** (2003), 1686–1692.
- [12] R. M. Neal, *Markov chain sampling methods for Dirichlet process mixture models*, Journal of Computational and Graphical Statistics **9** (2000), 249–265.
- [13] S. J. Pocock, N. L. Geller and A. A. Tsiatis, *The analysis of multiple endpoints in clinical trials*, Biometrics **43** (1987), 487–498.
- [14] A. E. Raftery and S. M. Lewis, *Implementing MCMC*, in: W. R. Gilks, D. J. Spiegelhalter and S. Richardson (eds.), Markov Chain Monte Carlo in Practice, Chapman and Hall, London, 1996, 115–130.
- [15] M. Sammel, X. Lin and L. Ryan, *Multivariate linear mixed models for multiple outcomes*, Statistics in Medicine **18** (1999), 2479–2492.
- [16] E. Strauss, E. Sherman and S. Otfried, *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*, 3rd ed., Oxford University Press, 2006.
- [17] J. L. Schafer, *Analysis of Incomplete Multivariate Data*, Monographs on Statistics & Applied Probability **72**, Chapman and Hall, London, 1997.
- [18] S. W. Thurston, D. Ruppert and P. W. Davidson, *Bayesian models for multiple outcomes nested in domains*, Biometrics **65** (2009), 1078–1086.
- [19] P. L. Williams, G. Molenberghs and S. R. Lipsitz, *Analysis of multiple ordinal outcomes in developmental toxicity studies*, Journal of Agricultural, Biological, and Environmental Statistics **1** (1996), 250–274.
- [20] L. Xiao, S. W. Thurston, D. Ruppert, T. M. Love and P. W. Davidson, *Bayesian models for multiple outcomes in domains with application to the Seychelles child development study*, Journal of the American Statistical Association **109** (2014), 1–10.

Amy LaLonde, Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY
e-mail: amy_lalonde@urmc.rochester.edu

Tanzy Love, Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY
e-mail: tanzy_love@urmc.rochester.edu