

Assessing Chemical Mixtures and Human Health: Use of Bayesian Belief Net Analysis

Anindya Roy¹, Neil J. Perkins², Germaine M. Buck Louis²

¹Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, USA; ²Division of Epidemiology, Statistics & Prevention Research, Eunice Kennedy Shriver National Institute of Child Health & Human Development, Rockville, USA.

Email: anindya@umbc.edu

Received April 18th, 2012; revised May 13th, 2012; accepted June 11th, 2012

ABSTRACT

Background: Despite humans being exposed to complex chemical mixtures, much of the available research continues to focus on a single compound or metabolite or a select subgroup of compounds inconsistent with the nature of human exposure. Uncertainty regarding how best to model chemical mixtures coupled with few analytic approaches remains a formidable challenge and served as the impetus for the study. **Objectives:** To identify the polychlorinated biphenyl (PCB) congener(s) within a chemical mixture that was most associated with an endometriosis diagnosis using novel graphical modeling techniques. **Methods:** Bayesian Belief Network (BBN) models were developed and empirically assessed in a cohort comprising 84 women aged 18 - 40 years who underwent a laparoscopy or laparotomy between 1999 and 2000; 79 (94%) women had serum concentrations for 68 PCB congeners quantified. Adjusted odds ratios (AOR) for endometriosis were estimated for individual PCB congeners using BBN models. **Results:** PCB congeners #114 (AOR = 3.01; 95% CI = 2.25, 3.77) and #136 (AOR = 1.79; 95% CI = 1.03, 2.55) were associated with an endometriosis diagnosis. Combinations of mixtures inclusive of PCB #114 were all associated with higher odds of endometriosis, underscoring its potential relation with endometriosis. **Conclusions:** BBN models identified PCB congener 114 as the most influential congener for the odds of an endometriosis diagnosis in the context of a 68 congener chemical mixture. BBN models offer investigators the opportunity to assess which compounds within a mixture may drive a human health effect.

Keywords: Bayesian Belief Network; Endometriosis; Environment; Mixtures; Polychlorinated Biphenyls

1. Introduction

Humans are exposed to diverse environmental chemical mixtures largely as a result of diet, occupation and modern life. However, much of the published literature focuses on a single compound or a select subset of compounds in relation to health outcomes, making it exceedingly difficult to synthesize the cumulative effects of chemical mixtures on human health, or to be sure the selected compound(s) is truly exerting the effect as presumed. As laboratory capabilities become even more sophisticated, allowing for the quantification of existing and emerging compounds at environmentally relevant concentrations, it is imperative to have analytical tools capable of modeling chemical mixtures in relation to health outcomes if we are to have a more complete understanding of the effects or modes of action.

Despite longstanding recognition of the importance of considering chemical mixtures, few analytic options are available to researchers when assessing possible health effects including those believed to adversely impact hu-

man reproduction and development as recently summarized [1,2]. Past analytic approaches for studying chemicals and endometriosis include: restricting to a particular compound such as dioxin [3], the *a priori* selection of more prevalent PCB congeners [4] or those with similar structural activity [5,6], the simple summing of individual congeners with or without a weighting as in the case of toxic equivalency quotients (TEQ) for dioxin [7], summing then categorizing by purported biologic activity such as estrogenic or anti-estrogenic congener groupings [8,9], or using differentially weighted linear combination of congeners to identify subsets that are most associated with the outcome [10].

The purported relation between exposure to hormonally active persistent environmental chemicals and endometriosis is an example of a suggestive albeit equivocal literature that has utilized varying analytic approaches for assessing this relation. For example, plasma dioxin was reported to be associated with endometriosis [3] while no association was observed using either a TEQ [7] or simple sum [4] approach for persistent organochlorine

compounds. However, when PCB congeners were categorized by purported biologic activity, an elevated odds of laparoscopically confirmed endometriosis was observed for anti-estrogenic PCB congeners [1]. While chemical profiles may vary across populations, it is plausible to assume that each population is exposed to a mixture of chemicals underscoring the need for investigators to attempt to utilize all available chemical profiling in assessing health effects to ensure our statistical models are indeed biologically plausible and reflective of contemporary geographic exposures.

We present an approach for assessing the effects of individual PCB congeners in the context of a mixture of congeners in relation to endometriosis, including congeners that may have opposing biologic activity, especially for an estrogenic dependent disease such as endometriosis [11]. The goal of our research is to utilize novel methods for identifying which, if any, compound within a chemical mixture is most associated with an adverse health effect such as endometriosis, thereby, allowing investigators to evaluate mixtures more closely reflecting human exposures.

2. Materials and Methods

2.1. Study Cohort and Data Collection

The study cohort comprised 84 (84% participation) women aged 18 - 40 years undergoing incident laparoscopy or laparotomy at one of two university-affiliated hospitals between 1999-2000. Standardized interviews were conducted primarily in the women's homes prior to surgery for the ascertainment of medical and reproductive history and current lifestyle. Upon completion of the interview, ≈ 20 cc of blood was obtained from 79 (94%) women using venipuncture equipment free of the contaminants under study. Surgeons were instructed to fully inspect the pelvis of all women regardless of preoperative diagnosis or symptoms, and to complete standardized operative reports regarding all observed gynecologic disorders. Thus, an endometriosis diagnosis was based upon laparoscopic visualization, which is considered the gold standard [12,13]. A more complete description of the study protocol is provided elsewhere [9].

2.2. Toxicological Analysis

Gas chromatography with electron-capture (GC-EC) was utilized for quantifying concentrations for 68 serum PCB congeners. Briefly, serum specimens were run in batches of 10 plus four quality control samples (*i.e.*, reagent blank, matrix blank, matrix blank containing a mixed standard of 15 specific congeners at known values, and a duplicate participant sample). Concentrations were calculated from standard curves for the 15 calibration standards; the re-

maining congener concentrations were calculated from response factors generated for all remaining laboratory measured congeners. Each congener concentration was adjusted for surrogate recovery and subtraction of reagent blanks [14]. The limit of detection was determined as three standard deviations from the mean of at least ten matrix blanks. Serum lipids were determined according to Phillips and colleagues' methods [15].

2.3. Statistical Analysis

In the descriptive phase, we first assessed the frequency distribution for each PCB congener to develop a categorization required for specifying the purported etiological models. There were 68 PCB congeners in the dataset, though six congeners (#6, #30, #47, #49, #50, #204) had only one nonzero observation (quantified in $<2\%$ of samples) and were deleted given that our statistical methodology required each congener to have at least three distinct observed percentiles. The remaining 62 congeners were used to build a joint probabilistic model for assessing the odds of an endometriosis diagnosis adjusting for *a priori* confounders. The goal was to establish a broad model that can answer various questions and possibly generate testable hypotheses about the effect of individual or groups of chemicals on human health outcomes such as endometriosis in the context of a chemical mixture.

Specifically, we used a Bayesian Belief Network (BBN) [16-19] to assess the impact of individual or groups of PCB congeners on the odds of an endometriosis diagnosis controlling for confounders. The BBN describes the joint probability distribution of a set of variables (*e.g.*, PCB congeners and endometriosis) via a series of conditional independence statements that are represented in the form of directed acyclic graphs. The graphs are representations of purported causal relations between exposures and an outcome [20]. Our specified BBN utilized a non-parametric formulation that eliminated the normality of data assumption [18], which is important given that many chemicals are not normally distributed. The BBN analysis was implemented using the Uninet 2.74 software [21], developed by Risk and Environmental Modeling Group at the Delft University of Technology, to estimate the model using empirical (conditional) rank correlations as computed from the data. Thus, the BBN was modeled to identify which, if any, congeners within the mixture were most associated with an elevated odds ratio for endometriosis.

A schematic diagram of the final BBN model is given in **Figure 1**. Each variable in the figure is shown as an elliptical node that may be influenced by other nodes via the directed arcs. The BBN in **Figure 1** has essentially three types of nodes: confounders (orange), PCB conge-

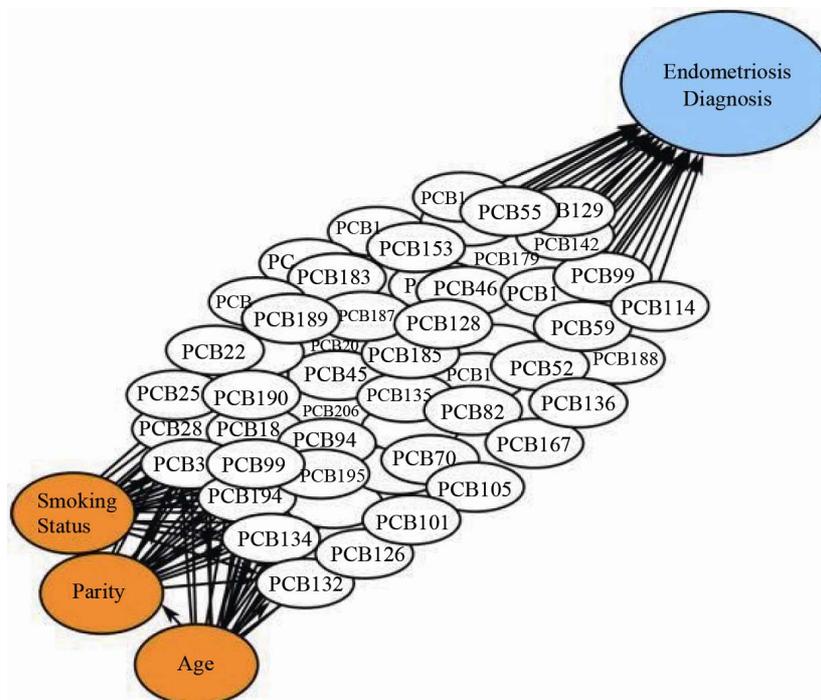


Figure 1. Proposed Bayesian belief network displaying the influence of 62 PCB congeners on endometriosis. The BBN is depicting the proposed causal structure for PCB congeners, confounders and endometriosis as designated by the white, orange and blue nodes, respectively.

ners (white) and the outcome or diagnosis of endometriosis (blue). The *a priori* confounders were: age (continuous), parity (nulliparous versus parous) and cigarette smoking status (yes/no) [9]. In the BBN, confounders may influence PCB congeners and the outcome, age may influence parity, and PCB congeners may influence the outcome. The BBN is used to estimate the adjusted odds ratios (AORs) and corresponding 95% confidence intervals (CIs) for endometriosis when the BBN is conditioned on specific concentration levels of individual PCB congeners or groups of congeners. Each conditional structure can be then interpreted as being representative of the population of women in the cohort with serum PCB concentrations approximately equal to the conditional values. Each arc is associated with weights representing the strength of the relation between the nodes joined by the arc. The weights can be set *a priori* by the user or can be empirically estimated from the data. Given the limited amount of information available about the effect of each PCB congener on endometriosis, we estimated the weights directly from the data using the default feature in the software. Given a set of weights, at each node one can then compute the conditional distributions of the current node given specific values of all other parent nodes. Once established, these distributions can then be used to evaluate a multitude of relations between PCB congeners, confounders and endometriosis within the specified framework. For simplicity purposes, we estimated AORs

and CIs for PCB congeners for a BBN conditioned on a relatively high value of the congener (75th percentile of the observed concentrations) and that for a low value (25th percentile of the observed concentrations).

3. Results

Thirty-two (38%) women were diagnosed with endometriosis, while 52 (62%) were not. As previously reported for this operative cohort, women diagnosed with endometriosis were older and more likely to be nulliparous and nonsmokers in comparison to unaffected women [9] underscoring the importance of controlling for these factors.

Use of the BBN identified PCB congener #114 as being highly associated with an elevated odds ratio of endometriosis even after adjusting for relevant covariates. Using the framework in **Figure 1**, the conditional structure led to a 0.24 probability of an endometriosis diagnosis for a woman with a 0.001 ng/g serum concentration of PCB #114, and a 0.48 probability of diagnosis for women with a higher concentration (0.013 ng/g serum). Specifically, the odds for endometriosis were about 3 times higher (AOR = 3.01; 95% CI = 2.25, 3.77) for women with elevated levels (75th percentile) of PCB #114 in comparison to women with relatively lower concentrations (the reference 25th percentile), representing a two times higher probability of an endometriosis diagnosis among women in the elevated group.

that focus either on an individual chemical or on a combination of selected congeners that have been simply summed and analyzed. We are unaware of any previous efforts attempting to analyze measured chemical concentrations, and more closely capturing individual exposure scenarios (irrespective of limits of detection), to explore potential mixtures in the context of other relevant biologic covariates and their effect on reproductive outcomes such as endometriosis. We utilized a BBN framework to explore the individual effects of PCB congeners and the odds of an endometriosis diagnosis so that the joint distribution of PCB concentrations and endometriosis could be assessed in the context of potential confounding among the measured concentrations. To this end, the BBN approach incorporates both a data-driven reduction approach to identify particular chemicals driving the effect conditional on all other exposures and biologically relevant covariates for endometriosis. Unlike other modeling approaches that may fail to reach convergence when sample size is limited relative to the number of chemicals being studied, the BBN approach is robust to this concern.

Our BBN identified a range of AORs for individual PCB congeners ranging from <1 to >3 . PCB #114 conferred the largest effect on the odds of an endometriosis diagnosis in this study cohort among the 62 PCB congeners assessed. In fact, high (75th percentile) levels of PCB #114 conferred three times the odds of endometriosis in comparison to low (25th percentile) levels. More interestingly, the effect of PCB #114 was greatly enhanced in mixtures with elevated levels of PCB congeners #99, #101, #136, and #153. In the context of the 75th percentile for PCB congeners #99, #101, #136, #153, the AOR for PCB #114 increased nearly 11-fold in comparison to women with all concentrations in the 25th percentile. Moreover, conditional on mixtures with high levels of PCB #114, the AOR for the other four influential congeners was 1.65. Even though the associated CI included one, the relatively large nominal value of the AOR points toward a possible incremental effect of the PCBs when evaluated in the context of a mixture. Combined these findings underscore the need to consider chemical mixtures, particularly since there may be geographic differences in the types of mixtures to which study populations are exposed. We believe that a more thorough analysis of all measured compounds (and the presentation of such data) is informative not only for the evaluation of potential health risks, but in delineating how mixtures may vary and whether an individual compound is etiologic irrespective of study population.

There are important limitations underlying our work that need to be considered when interpreting the results. Our intent was to demonstrate the feasibility of the BBN approach for assessing chemical mixtures to identify sig-

nals that may inform etiologic or mechanistic research. To this end, the BBN approach can be viewed as an empirically based data reduction approach and not one for determining etiology, *per se*. This is important given our cohort size, though it is comparable to many published studies on this topic. Other data-reduction methods such as principal component analysis (PCA), canonical correlation analysis, factor analysis and structural equation techniques often require normality assumptions for optimal performance. The utility of the BBN approach is its lack of assumptions required for parametric models including the normality assumption among others. Thus, findings from the BBN are more robust compared to those obtained from methods that rely on parametric assumptions.

Previous approaches that sum congeners essentially assign equal weights to all PCBs and, thereby, may not fully account for influential compounds. Gennings *et al.* [10] develop a novel method where differential weights are assigned to chemicals in a linear combination according to an optimization procedure, thereby allowing subsets of congeners to be differentially associated with the outcome. Methods such as PCA produce weighted sums of congeners on the basis of observed variability of the congeners in the mixture without accounting for specific of each congener on the outcome under study. Moreover, the weights in the linear combination need not be directly related to individual influences. In the BBN approach, the relative contribution of congeners in the mixture can be evaluated with regard to the analytic model as they are derived from the data resulting added flexibility. The BBN approach can serve as a tool for determining the relative influence of chemicals in a mixture when limited information is available to advise the investigator, and in generating testable hypotheses for future research.

Of all the PCB congeners assessed, #114 was most informative for endometriosis in this cohort of women. The extent to which #114 is etiologically associated with endometriosis awaits corroboration, but hopefully illustrates the utility of the BBN approach for identifying possible etiologic signals driving observed health effects for a particular cohort or study sample. Still, the biological interpretation of congeners identified as relevant to human health or disease outcomes through BBN approaches requires toxicologic and other biologic input. As with any modeling approach, our findings are dependent upon valid and reproducible laboratory measurements in the context of biologically plausible specified models. Clearly, a team science approach is needed in building the BBN and in interpreting the results. However, compounds identified as being important for disease outcomes may warrant subsequent experimental research aimed at determining mechanistic pathways to aid in the interpretation

of findings including possible statistical artifacts.

5. Conclusion

PCB congener #114 was identified by the BBN approach as the most influential compound within a mixture of 68 PCB congeners, conferring a 200% increased likelihood of an endometriosis diagnosis following laparoscopy. Thus, the BBN may be an approach for corroborating results across study populations, or in developing weighting schemes aimed at estimating the magnitude of effects exerted by individual compounds in keeping more closely with the manner in which human exposure arises.

6. Acknowledgements

The work was supported in part with funding from the American Chemistry Council and the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health & Human Development.

REFERENCES

- [1] G. M. Buck Louis, C. D. Lynch and M. A. Cooney, "Environmental Influences on Female Fecundity and Fertility," *Seminars in Reproductive Medicine*, Vol. 24, No. 3, 2006, pp. 147-155. [doi:10.1055/s-2006-944421](https://doi.org/10.1055/s-2006-944421)
- [2] G. Torf, L. Hagmar, A. Giwercman and J. P. Bonde, "Epidemiological Evidence on Reproductive Effects of Persistent Organochlorines in Humans," *Reproductive Toxicology*, Vol. 19, No. 1, 2004, pp. 5-26. [doi:10.1016/j.reprotox.2004.05.006](https://doi.org/10.1016/j.reprotox.2004.05.006)
- [3] A. Mayani, S. Barel, S. Soback and M. Almagor, "Dioxin Concentrations in Women with Endometriosis," *Human Reproduction*, Vol. 12, No. 2, 1997, pp. 373-375. [doi:10.1093/humrep/12.2.373](https://doi.org/10.1093/humrep/12.2.373)
- [4] B. Trabert, A. J. De Roos, S. M. Schwartz, U. Peters, D. Scholes, D. B. Barr, *et al.*, "Non-Dioxin-Like Polychlorinated Biphenyls and Risk of Endometriosis," *Environmental Health Perspectives*, Vol. 118, 2010, pp. 1280-1285. [doi:10.1289/ehp.0901444](https://doi.org/10.1289/ehp.0901444)
- [5] K. B. Moysich, P. Mendola, E. F. Schisterman, J. L. Freudenheim, C. B. Ambrosone, J. E. Vena, *et al.*, "An Evaluation of Proposed Frameworks for Grouping Polychlorinated Biphenyl (PCB) Congener Data into Meaningful Analytic Units," *American Journal of Industrial Medicine*, Vol. 35, No. 3, 1999, pp. 223-231. [doi:10.1002/\(SICI\)1097-0274\(199903\)35:3<223::AID-AJIM2>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-0274(199903)35:3<223::AID-AJIM2>3.0.CO;2-L)
- [6] M. S. Wolff, D. Camann, M. Gammon and S. D. Stellman, "Proposed PCB Congener Groupings for Epidemiological Studies," *Environmental Health Perspectives*, Vol. 105, No. 1, 1997, pp. 13-14. [doi:10.1289/ehp.9710513](https://doi.org/10.1289/ehp.9710513)
- [7] A. Pauwels, P. J. C. Schepens, T. D'Hooghe, L. Delbeke, M. Dhont, A. Brouwer, *et al.*, "The Risk of Endometriosis and Exposure to Dioxins and Polychlorinated Biphenyls: A Case-Control Study of Infertile Women," *Human Reproduction*, Vol. 16, No. 10, 2001, pp. 2050-2055. [doi:10.1093/humrep/16.10.2050](https://doi.org/10.1093/humrep/16.10.2050)
- [8] P. S. Cooke, T. Sato and D. L. Buchanan, "Disruption of Steroid Hormone Signaling by PCBs," In: L. W. Robertson and L. G. Hansen, Eds., *PCBs: Recent Advances in Environmental Toxicology and Health Effects*, The University Press of Kentucky, Lexington, 2001, pp. 257-263.
- [9] G. M. Buck Louis, J. M. Weiner, B. W. Whitcomb, R. Sperrazza, E. F. Schisterman, D. T. Lobbell, *et al.*, "Environmental PCB Exposure and Risk of Endometriosis," *Human Reproduction*, Vol. 20, No. 1, 2005, pp. 279-285. [doi:10.1093/humrep/deh575](https://doi.org/10.1093/humrep/deh575)
- [10] C. Gennings, R. Sabo and E. Carney, "Identifying Subsets of Complex Mixtures Most Associated with Complex Diseases: Polychlorinated Biphenyls and Endometriosis as a Case Study," *Epidemiology*, Vol. 21, 2010, pp. 77-84. [doi:10.1097/EDE.0b013e3181ce946c](https://doi.org/10.1097/EDE.0b013e3181ce946c)
- [11] G. S. DiZerega, D. L. Barber and G. D. Hodgen, "Endometriosis: Role of Ovarian Steroids in Initiation, Maintenance, and Suppression," *Fertility and Sterility*, Vol. 33, 1980, pp. 649-653.
- [12] S. Kennedy, A. Bergqvist, C. Chapron, T. D'Hooghe, G. Dunselmans, R. Greb, *et al.*, "ESHRE Guidelines for the Diagnosis and Treatment of Endometriosis," *Human Reproduction*, Vol. 20, No. 10, 2005, pp. 2698-2704. [doi:10.1093/humrep/dei135](https://doi.org/10.1093/humrep/dei135)
- [13] American Society for Reproductive Medicine, Practice Committee, "Endometriosis and Infertility," *Fertility and Sterility*, Vol. 86, No. 5, 2006, pp. S156-S160. [doi:10.1016/j.fertnstert.2006.08.014](https://doi.org/10.1016/j.fertnstert.2006.08.014)
- [14] M. S. Bloom, G. M. Buck Louis, E. F. Schisterman and A. Liu, "Maternal Serum Polychlorinated Biphenyl Concentrations across Critical Windows of Human Development," *Environmental Health Perspectives*, Vol. 115, 2007, pp. 1320-1324. [doi:10.1289/ehp.10086](https://doi.org/10.1289/ehp.10086)
- [15] D. J. Phillips, J. L. Pirkle, V. W. Burse, J. T. Pernert, L. O. Henderson and L. L. Needham, "Chlorinated Hydrocarbon Levels in Human Serum: Effects of Fasting and Feeding," *Archives of Environmental Contamination and Toxicology*, Vol. 18, No. 4, 1989, pp. 495-500. [doi:10.1007/BF01055015](https://doi.org/10.1007/BF01055015)
- [16] T. J. Bedford and R. M. Cooke, "Vines—A New Graphical Model for Dependent Random Variables," *Annals of Statistics*, Vol. 30, No. 4, 2002, pp. 1031-1068. [doi:10.1214/aos/1031689016](https://doi.org/10.1214/aos/1031689016)
- [17] A. M. Hanea, D. Kurowicka and R. M. Cooke, "Hybrid Method for Quantifying and Analyzing Bayesian Belief Nets," *Quality and Reliability Engineering International*, Vol. 22, No. 6, 2006, pp. 709-729. [doi:10.1002/qre.808](https://doi.org/10.1002/qre.808)
- [18] A. M. Hanea, D. Kurowicka, R. M. Cook and D. M. Ababei, "Mining and Visualizing Ordinal Data with Non-Parametric Continuous BBNs," *Computational Statistics & Data Analysis*, Vol. 54, No. 3, 2010, pp. 668-687. [doi:10.1016/j.csda.2008.09.032](https://doi.org/10.1016/j.csda.2008.09.032)
- [19] D. Kurowicka and R. M. Cooke, "Uncertainty Analysis and High Dimensional Dependence Modeling," Wiley, New York, 2006. [doi:10.1002/0470863072](https://doi.org/10.1002/0470863072)
- [20] S. Greenland, J. Pearl and J. M. Robins, "Causal Dia-

grams for Epidemiologic Research,” *Epidemiology*, Vol. 10, No. 1, 1999, pp. 37-48.
[doi:10.1097/00001648-199901000-00008](https://doi.org/10.1097/00001648-199901000-00008)

[21] Risk and Environmental Modeling Group, “Uninet 2.74 Software,” Delft University of Technology, 2008.
<http://dutiosc.twi.tudelft.nl/~risk/>