Environmental Research 160 (2018) 115–120

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Latent class analysis to model multiple chemical exposures among children

Michael Hendryx[a,*], Juhua Luo[b]

[a] Department of Environmental and Occupational Health, School of Public Health, Indiana University, 1025 E 7th St., Bloomington, IN 47405, United States
[b] Department of Epidemiology and Biostatistics, School of Public Health, Indiana University, 1025 E 7th St., Bloomington, IN 47405, United States

ARTICLE INFO

Keywords:
Children
NHANES
Latent class analysis
Chemical exposures
Race and ethnicity

ABSTRACT

Background: Children are exposed to multiple potentially harmful chemicals simultaneously. Efforts to understand the patterns and consequences of these exposures have been hampered by statistical limitations in estimations of higher order interactions.

Objectives: The current study uses latent class analysis, a form of person-centered modeling to identify unobservable subgroups within populations and examine relationships between latent classes and measures of immune function.

Methods: Data from the National Health and Nutrition Examination Survey 2011–2012 were analyzed. A sample of 721 children aged 6–19 years were included who provided data on 47 chemicals of interest representing six chemical classes. Groups were identified using latent class analysis controlling for race/ethnicity, age, sex and poverty status.

Results: Two alternative approaches to identifying latent classes each resulted in similar three class solutions, including one group of children characterized by low co-exposures across chemicals, a group with moderate co-exposure levels, and a group characterized by high co-occurring levels of polycyclic aromatic hydrocarbons, volatile organic compounds, phenols and phthalates. Under one of the approaches, latent classes were significantly associated with immune function as measured by lymphocyte and neutrophil counts.

Conclusions: Latent class analysis offers a potential approach to measuring and understanding interactions among multiple co-occurring chemical stressors. However, additional work is needed to test the ability of latent classes to predict health variables.

Exposures to multiple chemicals simultaneously is widespread among adults (CDC, 2009; Rosofsky et al., 2016; Woodruff et al., 2011) and children (Calafat et al., 2016; Zota et al., 2014). Children are more susceptible to the effects of chemical exposures that adults (ATSDR, 2013; NRC, 1993; Perera, 2017). This is because children are experiencing sensitive developmental windows characterized by rapid growth and maturation of cells and organ systems. Chemical insults during these sensitive periods may have more profound and long term impacts than similar exposures among adults. Children are also more susceptible than adults because of their small size, so that a given exposure represents a larger relative dose. Children also engage in different behaviors that may place them at greater exposure risk (ATSDR, 2013); for example, they may play close to carpets or floor surfaces that contain various chemicals, they may place contaminated objects in their mouths, or they may come into close contact with soil or plants treated with harmful chemicals.

Increasing awareness has been expressed within the scientific community about the potential additive or synergistic impacts that children may experience from multiple co-occurring exposures (Calafat et al., 2016; Choi et al., 2017; NIH, 2016; Steer et al., 2015). However, research to understand the patterns and consequences of multiple co-occurring chemical exposures is underdeveloped. Efforts to study co-occurring exposures have usually been limited to a few, even two, such exposures at a time (Chen et al., 2014; Hambach et al., 2013). In contrast, the concerns about multiple exposures as expressed through such constructs as the total environment framework (Tulve et al., 2016), cumulative risk assessment (Wason et al., 2012; Williams et al., 2012) or the related concept of the exposome (Buck Louis et al., 2017; Manrai et al., 2017; Patel, 2017; Wild, 2012) demand that we consider multiple simultaneous influences, certainly more than two at a time.

The problem of understanding multiple co-occurring exposures is complicated through the limitations imposed by conventional data analysis approaches. Conventional approaches are variable centered, meaning that correlations among sets of variables are examined within a sample. Understanding the synergistic effects of multiple variables within a sample requires estimating higher order interactions.

* Correspondence to: School of Public Health, Indiana University, 1025 E 7th St., Bloomington, IN 47405, United States.
E-mail address: hendryx@indiana.edu (M. Hendryx).

http://dx.doi.org/10.1016/j.envres.2017.09.024
Received 2 June 2017; Received in revised form 19 September 2017; Accepted 21 September 2017

Estimating even two-way interactions can impose restrictions on statistical power, a problem which is magnified as interaction orders become greater. Furthermore, interpretation of higher order interactions, even if they can be estimated, is unwieldly. This statistical limitation accounts in part for the few studies that have examined multiple co-occurring exposures.

In contrast to variable centered models, an alternative approach is to estimate person-centered models. Specifically, we use latent class analysis, a form of person-centered modeling that identifies unobservable subgroups within populations who experience unique patterns of multiple risks (Lanza et al., 2007). The method is designed to identify higher order interactions among multiple stressors and overcomes the statistical power limitations of identifying these interactions using variable-centered approaches. Latent class analysis has been used primarily within the social sciences to model multiple behavioral risks (Hillemeier et al., 2013; Lanza et al., 2011). Ko et al. (Ko et al., 2016) illustrated its application to understanding dietary patterns. To date, however, it has not been applied to the problem of multiple chemical stressors.

In the current study we illustrate an application of this method using National Health and Nutrition Examination (NHANES) data. Ultimately each person is assigned a probability of belonging to each of a number of limited latent classes that are characterized by unique combinations of multiple exposures. Our intent is to offer a novel approach for identification of interactions that may overcome the limitations of conventional statistical analytic attempts. We will attempt to demonstrate 1) that there are a discrete number of latent classes that characterize children's exposures to multiple chemical groups, 2) that these latent classes will be related to selected health status indicators, in this case to immune function as measured by counts of lymphocytes and neutrophils (Jacobs et al., 2010; Kimber and Dearman, 2010; Sarasua et al., 2000), and 3) that latent classes that capture multiple co-exposures will offer a distinctly unique interpretation of co-exposure effects that traditional main effects models cannot.

1. Methods

1.1. Design

The study is a non-experimental analysis of National Health and Nutrition Examination (NHANES) data for the 2011–2012 cycle. This cycle was selected because the same children were simultaneously measured across multiple chemical classes of interest. The NHANES is a biennial, nationally representative sample of non-institutionalized US residents conducted by the Centers for Disease Control and Prevention (CDC). Additional detail regarding sampling procedures and analytic methods may be found from the CDC NHANES site (CDC, 2016).

1.2. Sample

We included 721 children ages 6–19 years with data on all variables of interest for analysis. The chemicals of interest were not collected from children less than 6 years of age.

1.3. Measures

We selected chemical exposure measures that were available among the same children, that are potentially of health concern, and that were usually present above detection levels. We initially considered 62 such chemicals, all measured via urinary analytes or metabolites, representing six chemical classes: (11 metals, 2 pesticides, 3 phenols, 11 phthalates, 25 volatile organic compounds (VOCs), and 10 polycyclic aromatic hydrocarbons (PAHs)). Levels of chemicals were adjusted for creatinine. All chemicals were measured as μ g/L except PAHs and VOCs, measured as ng/L.

Rather than trying to model all 62 chemicals, we attempted to identify a parsimonious set of chemicals using two alternative approaches. In the first approach, we identified a set of chemicals within each chemical group that maximized the internal consistency reliability of the within-group correlations. To do this we found the Cronbach's alpha reliability scores for each group, and eliminated individual chemicals that reduced the reliability of those scores. Most within-class chemical correlations were positive and significant, and we retained 47 of the original 62 chemicals. We kept all 10 PAHs and both of the original pesticides. We retained 9 of 11 metals, 17 of 25 VOCs, 2 of 3 phenols, and 7 of 11 phthalates. Then, we calculated the standardized score of each retained chemical, setting a mean of 100 and standard deviation of 10, and found the mean score for each of the standardized groups. The standardization allows us to weight each chemical group equally and set them on common scales. Table 1 lists the 47 retained chemicals using names provided in the NHANES database. The final standardized Cronbach alpha values for the six groups were: metals (0.97), VOCs (0.99), pesticides (0.78), phenols (0.47), phthalates (0.89), and PAHs (0.99).

In the second approach, we simply choose the single chemical from each of the six chemical groups that had the highest correlation to the total chemical group. These chemicals are marked with the asterisks in Table 1. Both the standardization approach and the single chemical approach resulted in six items for latent class modeling

In addition to the chemical measures, we included as descriptive variables and covariates: age in years (divided into younger (aged 6–12) and older (aged 13–19)), sex, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and Other (including Asian, multiracial and other groups)), and poverty status (yes/no based on

Table 1
List of chemicals used in latent class analysis.

Metals	**Phthalates**
Antimony	Mono − 2-ethyl − 5-carboxypentyl
Barium	Mono-n-butyl
Cadmium	Mono-(3-carboxypropyl)
Cesium	Mono-(2-ethyl − 5-hydroxyhexyl)
Cobalt	Mono-(2-ethyl − 5-oxohexyl)[a]
Lead	Mono-isobutyl
Thallium[a]	Mono-benzyl
Tungsten	**Volatile Organic Compounds (VOCs)**
Uranium	N-Acetyl-S-(1,2-dichlorovinyl)-L-cysteine
Pesticides	N-Acetyl-S-(2,2-dichlorovinyl)-L-cysteine
2,5-dichlorophenol[a]	2-Methylhippuric acid
2,4-dichlorophenol	3- and 4-Methylhippuric acid
Polycyclic aromatic hydrocarbons (PAHs)	N-Acetyl-S-(2-carbamoylethyl)-L-cysteine
1-hydroxynaphthalene	N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine
2-hydroxynaphthalene	2-Aminothiazoline − 4-carboxylic acid
2-hydroxyfluorene	N-Acetyl-S-(benzyl)-L-cysteine
3-hydroxyfluorene	N-Acetyl-S-(2-carboxyethyl)-L-cysteine
1-hydroxyphenanthrene[a]	N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine[a]
2-hydroxyphenanthrene	N-Acetyl-S-(dimethylphenyl)-L-cysteine
3-hydroxyphenanthrene	N-Acetyl-S-(2-carbamoyl − 2-hydroxyethyl)-L-cysteine
1-hydroxypyrene	N-Acetyl-S-(2-hydroxypropyl)-L-cysteine
9-hydroxyfluorene	N-Acetyl-S-(3-hydroxypropyl)-L-cysteine
4-phenanthrene	Mandelic acid
Phenols	Phenylglyoxylic acid
Benzophenone − 3[a]	N-Acetyl-S-(3-hydroxypropyl − 1-methyl)-L-cysteine
Triclosan	N-Acetyl-S-(trichlorovinyl)-L-cysteine
	2-Thioxothiazolidine − 4-carboxylic acid

[a] Selected for the model containing one chemical per group.