

UCSF

UC San Francisco Previously Published Works

Title

Subpopulations of children with multiple chronic health outcomes in relation to chemical exposures in the ECHO-PATHWAYS consortium.

Permalink

<https://escholarship.org/uc/item/1jx181xg>

Authors

Day, Drew

LeWinn, Kaja

Karr, Catherine

et al.

Publication Date

2024-03-01

DOI

10.1016/j.envint.2024.108486

Peer reviewed



Published in final edited form as:

Environ Int. 2024 March ; 185: 108486. doi:10.1016/j.envint.2024.108486.

Subpopulations of children with multiple chronic health outcomes in relation to chemical exposures in the ECHO-PATHWAYS consortium

Drew B. Day^{a,*}, Kaja Z. LeWinn^b, Catherine J. Karr^{c,d,e}, Christine T. Loftus^c, Kecia N. Carroll^{f,g}, Nicole R. Bush^{b,h}, Qi Zhaoⁱ, Emily S. Barrett^{j,k}, Shanna H. Swan^f, Ruby H.N. Nguyen^l, Leonardo Trasande^m, Paul E. Mooreⁿ, Ako Adams Ako^o, Nan Ji^p, Chang Liu^q, Adam A. Szpiro^r, Sheela Sathyanarayana^{a,c,d,e} **Environmental influences on Child Health Outcomes**

^aCenter for Child Health, Behavior, and Development, Seattle Children's Research Institute, 1920 Terry Avenue, Seattle, Washington 98101, USA

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author at: M/S Cure-3, PO Box 5371, Seattle, WA 98145. drew.day@seattlechildrens.org (D.B. Day).

Additional Contributions

We would like to acknowledge the contributions of the following staff members across our multiple sites and cohorts. As part of the CANDLE study, the following individuals contributed to data collection: Maureen Sorrells, April Huggins, Regina Werkhoven, Anum Abbas, Ryan Carr, L. Ashley Robinson, and Morgan Bromley. The following individuals contributed to data collection as part of the TIDES study: Stephanie Grover, Simar Singh, Beomyun Han, Zainab Ghadialy, Abigail Gelineau, Lauren Berg, Pamela Carr-Manthe, Alexandra Gowdy-Jaehnig, Andrea Hart, Annabel Victor-Halliday, Tye Johnson, and Thomson Kelly. The following individuals contributed to data collection as part of the GAPPS-PW study: Maria Hernandez, Daisy Hernandez, Madison Grady, Ifeoma Okafor, Kendall Dunlop Korsness, Janine Mifsud, Diane Bowman, and Mehar Maju. Logan Dearborn, Erin Buth, and Elizabeth Spalt contributed to data management. All non-author contributors gave permission to be included.

CRedit authorship contribution statement

Drew B. Day: Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Kaja Z. LeWinn:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Catherine J. Karr:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Christine T. Loftus:** Investigation, Methodology, Writing – review & editing. **Kecia N. Carroll:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Nicole R. Bush:** Funding acquisition, Investigation, Methodology, Project administration, Resources, Writing – review & editing. **Qi Zhao:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Emily S. Barrett:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Shanna H. Swan:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Ruby H.N. Nguyen:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Leonardo Trasande:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Paul E. Moore:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Ako Adams Ako:** Writing – review & editing. **Nan Ji:** Writing – review & editing. **Chang Liu:** Writing – review & editing. **Adam A. Szpiro:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Sheela Sathyanarayana:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: All authors reports financial support was provided by National Institutes of Health. Drew B. Day & Emily S. Barrett reports financial support was provided by National Institute of Environmental Health Sciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Conflict of Interest Disclosures

Drs. Day, LeWinn, Karr, Loftus, Carroll, Bush, Zhao, Barrett, Swan, Nguyen, Trasande, Moore, Ako, Ji, Liu, Szpiro, and Sathyanarayana reported receiving grants from the National Institutes of Health during the conduct of this study. Dr. Day and Dr. Barrett reported receiving grants from the National Institute of Environmental Health Sciences during the conduct of the study. No other disclosures were reported

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2024.108486>.

^bDepartment of Psychiatry and Behavioral Sciences, University of California, San Francisco, 675 18th Street, San Francisco, CA 94143, USA

^cDepartment of Environmental and Occupational Health, University of Washington, 4245 Roosevelt Way NE, Seattle, WA 98105, USA

^dDepartment of Epidemiology, University of Washington, 4245 Roosevelt Way NE, Seattle, WA 98105, USA

^eDepartment of Pediatrics, University of Washington, 4245 Roosevelt Way NE, Seattle, WA 98105, USA

^fDepartment of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Place, New York, NY 10029, USA

^gDepartment of Pediatrics, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Place, New York, NY 10029, USA

^hDepartment of Pediatrics, University of California San Francisco, San Francisco, CA, USA

ⁱDepartment of Preventive Medicine, Division of Preventive Medicine, University of Tennessee Health Science Center, 66 North Pauline Street, Memphis, TN 38163, USA

^jDepartment of Biostatistics and Epidemiology, Rutgers School of Public Health, 683 Hoes Lane West, Piscataway, NJ 08854, USA

^kEnvironmental and Occupational Health Sciences Institute, Rutgers University, 170 Frelinghuysen Road, Piscataway, NJ 08854, USA

^lDepartment of Epidemiology and Community Health, University of Minnesota, 420 Delaware Street Southeast, Minneapolis, Minnesota 55455, USA

^mDepartment of Pediatrics, New York University Grossman School of Medicine, 550 First Avenue, New York, NY 10016, USA

ⁿDivision of Allergy, Immunology, and Pulmonary Medicine, Department of Pediatrics, Vanderbilt University Medical Center, 2200 Children's Way, Nashville, TN 37232, USA

^oDepartment of Pediatrics, Children's Hospital at Montefiore, 3415 Bainbridge Avenue, Bronx, NY 10467, USA

^pDivision of Environmental Health, Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, 1845 N Soto St, MC 9239, Los Angeles, CA, 90039, USA

^qDepartment of Psychology, Washington State University, Johnson Tower, Pullman, WA 99164, USA

^rDepartment of Biostatistics, University of Washington, 3980 15th Avenue NE, Seattle, WA 98195, USA

Abstract

A multimorbidity-focused approach may reflect common etiologic mechanisms and lead to better targeting of etiologic agents for broadly impactful public health interventions. Our aim was

to identify clusters of chronic obesity-related, neurodevelopmental, and respiratory outcomes in children, and to examine associations between cluster membership and widely prevalent chemical exposures to demonstrate our epidemiologic approach. Early to middle childhood outcome data collected 2011–2022 for 1092 children were harmonized across the ECHO-PATHWAYS consortium of 3 prospective pregnancy cohorts in six U.S. cities. 15 outcomes included age 4–9 BMI, cognitive and behavioral assessment scores, speech problems, and learning disabilities, asthma, wheeze, and rhinitis. To form generalizable clusters across study sites, we performed k-means clustering on scaled residuals of each variable regressed on study site. Outcomes and demographic variables were summarized between resulting clusters. Logistic weighted quantile sum regressions with permutation test p-values associated odds of cluster membership with a mixture of 15 prenatal urinary phthalate metabolites in full-sample and sex-stratified models. Three clusters emerged, including a healthier Cluster 1 (n = 734) with low morbidity across outcomes; Cluster 2 (n = 192) with low IQ and higher levels of all outcomes, especially 0.4–1.8-standard deviation higher mean neurobehavioral outcomes; and Cluster 3 (n = 179) with the highest asthma (92 %), wheeze (53 %), and rhinitis (57 %) frequencies. We observed a significant positive, male-specific stratified association (odds ratio = 1.6; p = 0.01) between a phthalate mixture with high weights for MEP and MHPP and odds of membership in Cluster 3 versus Cluster 1. These results identified subpopulations of children with co-occurring elevated levels of BMI, neurodevelopmental, and respiratory outcomes that may reflect shared etiologic pathways. The observed association between phthalates and respiratory outcome cluster membership could inform policy efforts towards children with respiratory disease. Similar cluster-based epidemiology may identify environmental factors that impact multi-outcome prevalence and efficiently direct public policy efforts.

Keywords

Clustering; Multi-outcome; Multi-morbidity; Phthalates; Asthma; Behavior

1. Introduction

Approximately one in four children have a chronic medical condition in the United States, and up to one in five children have a psychiatric diagnosis at some point during childhood (Perou et al., 2013; Van Cleave et al., 2010). Mental health disorders, asthma and other allergic respiratory conditions, and obesity are among the most prevalent of these chronic conditions (Perrin et al., 2007). Comorbidity may across disease domains, such as 8–16 % of pediatric asthma cases being comorbid with obesity or psychiatric disorders (Kaplan et al., 2020). Complex multimorbidity patterns have been revealed by latent class analysis (Carrilero et al., 2020). Multi-cohort prospective studies can be a valuable resource in understanding factors underlying chronic comorbidity, as they combine rich exposure assessment with large quantities of outcome data.

Examining exposure associations with multi-outcome phenotypes may help to identify common etiologic mechanisms underlying multiple disease domains, such as asthma and depression sharing common deficits in adrenergic and glucocorticoid receptor expression (Miller and Chen, 2006). This approach may also lead to more effective

public health strategies by enabling identification of exposures that impact multiple diseases with high public health importance (VanderWeele, 2017). This is of particular relevance to environmental epidemiology since many environmental stressors share common pathophysiologic mechanisms with broad biologic implications such as endocrine disruption and oxidative stress-induced inflammation (Neier et al., 2015). Phthalates, which are ubiquitous synthetic endocrine-disrupting chemicals that can cross the placenta (Hauser and Calafat, 2005; Silva et al., 2004), represent an excellent case study in the potential utility of this outcome-wide approach since they have previously been associated, often in a sex-specific manner (Eales et al., 2022), with obesity (Kim and Park, 2014), asthma-related respiratory outcomes (Li et al., 2017), and both cognitive and behavioral neurodevelopmental outcomes (Radke et al., 2020).

Our goal was to identify and characterize clusters of children with correlated neurodevelopmental, respiratory, and obesity-related outcomes collected across multiple observational cohort studies. For this analysis, we utilized the ECHO-PATHWAYS consortium, which compiles a rich set of neurobehavioral, respiratory, and anthropometric outcomes from three separate cohorts and a total of 6 study sites across the U.S. (LeWinn et al., 2022). We applied a clustering method capable of disentangling the clustering structure induced by spatial, temporal, and technical differences between study cohorts and sites in mixed-type continuous and categorical data. To demonstrate how these multi-outcome phenotypes may be applied to environmental epidemiology, we additionally examined associations between prenatal phthalate exposures and the probability of membership in clusters of multimorbid outcomes. Our hypothesis is that we will discover stable clusters of multiple health outcomes, and that there will be potentially sex-specific associations between the probability of belonging to these clusters and prenatal exposure to mixtures of phthalates.

2. Materials and methods

2.1. Study population

This study utilized data on participants within the ECHO-PATHWAYS Consortium, which is a consortium within the larger Environmental Influences on Child Health Outcomes (ECHO) Program (Knapp et al., 2023) that includes three cohorts: The Conditions Affecting Neurocognitive Development and Learning in Early childhood (CANDLE) study, the Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) study, and The Infant Development and the Environment Study (TIDES) (LeWinn et al., 2022). The CANDLE study is a prospective pregnancy cohort in Shelby County, Tennessee that recruited pregnant women between 16 and 40 years old from 2006 to 2011, and the study design and inclusion criteria have been described elsewhere (Sontag-Padilla et al., 2015). Women attended two prenatal study visits and one at delivery that included maternal surveys and biospecimen collection. Mother-child dyads were followed up at regular intervals, including clinic visits at age 4–6 years and 8–9 years. The GAPPS study prospectively enrolled pregnant women over 18 years old over 8 years from hospitals in Seattle and Yakima, WA (LeWinn et al., 2022). Biospecimen collection occurred over three visits during pregnancy. Women previously enrolled in GAPPS with children born between 2011 and 2016 were re-contacted and

invited to participate in the ECHO-PATHWAYS Consortium. Visits including clinic and remote visits were performed for mother–child dyads at age 4–6 years upon enrollment into ECHO-PATHWAYS and later at age 8–9 years. TIDES is a multi-site prospective cohort that recruited pregnant women over 18 years old with low-risk pregnancies from 2010 to 2012 at four obstetrical clinics affiliated with participating academic medical institutions across the US: San Francisco, CA; Minneapolis, MN; Rochester, NY; and Seattle, WA (Barrett et al., 2014). Three prenatal study visits including biospecimen and survey collection were performed, and postnatal follow-up visits included a birth exam and clinic and remote visits at ages 4–5, 6–8, and 8–9 years.

2.2. Outcomes

We included 15 outcomes assessed during early to middle childhood representing the broad categories of obesity-related, respiratory, and neurodevelopmental outcomes. To represent obesity-related outcomes, we included body mass index (BMI) z-scores adjusted for age and sex assigned at birth using CDC reference values (Kuczmarski et al., 2002; van Buuren, 2018). Height and weight measurements used in this calculation were measured primarily during age 8–9-year visits. When these measurements were unavailable, we instead used age- and sex-adjusted BMI z-scores derived from age 4–6 visits (CANDLE: 2.6 % and GAPPS: 21 %) or an age 6–8 visit (TIDES: 26 %). Pearson correlation coefficients between repeated BMI z-scores from these visits were 0.66 for CANDLE, 0.82 for GAPPS, and 0.76 for TIDES.

Neurodevelopmental outcomes included cognitive performance indicated by full-scale intelligence quotient (FSIQ), harmonized as previously described (Ni et al., 2022) across age 4–6 Stanford Binet 5 (Roid and Barram, 2004); age 4–6 Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (Wechsler, 2012); and age 6–8 Wechsler Intelligence Scale for Children, Fifth Edition (Wechsler, 2014) assessments from CANDLE, GAPPS, and TIDES, respectively. Behavioral outcomes included age 8–9 anxiety, depression, and irritability as assessed using the total scores for the child self-reported Screen for Child Anxiety Related Emotional Disorders (SCARED) assessment (Birmaher et al., 1997), Children’s Depression Inventory 2 (CDI) (Kovacs, 1992), and the Affective Reactivity Index (ARI) (Stringaris et al., 2012), respectively. We also included five Child Behavior Checklist (CBCL) (Achenbach, 2001) subscale scores of social, thought, and attention problems as well as externalizing and internalizing behaviors reported by parents at age 8–9. Health history questionnaire-reported diagnosis of child learning disability or speech problems as reported by parents at age 4–6 for GAPPS and TIDES and 8–9 for CANDLE were also included among the neurodevelopmental outcomes.

Finally, age 8–9 parent-reported International Study of Asthma and Allergies in Childhood (ISAAC) (Asher et al., 1995) questionnaire-derived ever asthma, wheeze in the past 12 months, and ever allergic rhinitis or hay fever were included as respiratory outcomes. Ever wheeze was not used since this is less specific to chronic symptoms and may be related to bronchiolitis resulting from common early life infections (Smyth and Openshaw, 2006). The questions used for each dichotomous outcome are listed in the Supplementary Methods.

2.3. Maternal urine collection and phthalate metabolite measurement

Single maternal spot urine samples were collected during a late pregnancy visit with a median gestational age at collection of 32 weeks. Detailed urine collection methods for CANDLE (Sontag-Padilla et al., 2015), GAPPs (Wallace et al., 2023), and TIDES (Swan et al., 2015) have been described previously. 22 monoester phthalate metabolites (Table S5) were measured using previously described methods (Guo et al., 2014; Rocha et al., 2017) at the Wadsworth Laboratory at the New York State Department of Health, a member of the Children's Health Exposure Analysis Resource laboratory network (Kannan et al., 2021). Specific gravity was measured upon urine collection using a handheld refractometer. Phthalates with > 60 % detection were included, and undetected values were imputed with the limit of detection/ $\sqrt{2}$ (Hornung and Reed, 1990). The molar sum of five di(2-ethylhexyl) phthalate (DEHP) metabolites was calculated for analysis in the individual linear regressions (Swan et al., 2015).

2.4. Covariates

For the regression analyses associating prenatal urinary phthalate concentrations with multi-outcome clusters, we adjusted for a rich set of covariates based on previous studies associating prenatal phthalates with subsets of the outcomes included in our analysis. These included study site, child sex assigned at birth, maternal age at delivery, \log_2 -transformed time- and site-adjusted household income, categorical household count, maternal pre-pregnancy BMI, dichotomous primiparity, dichotomous maternal ever asthma diagnosis, dichotomous maternal prenatal smoking, maternal race, dichotomous Hispanic ethnicity, maternal educational attainment, maternal stress score as derived from the Perceived Stress Scale (PSS-4) (Cohen et al., 1983) collected as previously described (Wallace et al., 2023), and year of birth. Year of birth was included to control for temporal trends in outcomes and phthalate exposures (Zota et al., 2014). Race and ethnicity were included in models as proxies for systemic social and health inequities that may impact both phthalate exposure (Attina et al., 2019) as well as obesity-related (Isong et al., 2018), respiratory (Davis et al., 2021), and neurodevelopmental (Alegria et al., 2022) health outcomes. Additional urinary analysis-related covariates controlled in the models to limit exposure misclassification bias were urinary specific gravity, laboratory analysis batch, and gestational age at sample collection. Maternal prenatal smoking was defined as either self-reported smoking during pregnancy or urinary cotinine levels > 200 ng/mL (Schick et al., 2017). A directed acyclic graph showing proposed relationships between all variables is shown in Figure S6.

2.5. Statistical methods

In order to determine the optimal method for identifying latent multi-outcome disease profiles independent of between-site differences, we tested several methods for identifying the latent disease profile cluster memberships (Supplementary Methods). Briefly, we tested the identification of simulated latent disease clusters by evaluating the adjusted Rand index (ARI) (Hubert and Arabie, 1985) values across 500 iterations for several forms of agglomerative clustering and latent class model algorithms, altering the degree of separation (Qiu and Joe, 2006) between clusters, numbers of observations and clusters, noise variables, and other parameters. The method that performed best under almost every condition was one

we called residual k-means (Supplementary Methods), which involves separately regressing each variable on study site in a univariate linear regression and then performing k-means clustering on the centered and scaled residuals (Hartigan and Wong, 1979). Specifically, these were standard linear regression residuals for the continuous outcomes and response residuals (i.e., each 0 or 1 outcome value subtracted by the fitted values) from logistic regressions for categorical outcomes. We applied residual k-means to cluster complete outcome data. The optimal number of partitions were chosen as the value from 2 to 10 most commonly chosen among 24 different default algorithms implemented in the *NbClust* R package (Charrad et al., 2014) using the “complete” method. Cluster stability was evaluated with the average Jaccard index of 1000 bootstrap k-means iterations (Hennig, 2007).

Individual associations between \log_{10} -transformed prenatal urinary phthalates and cluster membership were assessed in logistic regressions omitting non-comparison clusters for the equivalent of multinomial regression (Hennig, 2007). Models were controlled for all aforementioned covariates first without and then with examination of effect modification by sex, which was implemented by including multiplicative interaction terms between sex and all independent variables in each model (Buckley et al., 2017). Associations between mixtures of 15 specific gravity-adjusted (Levine, 1945) phthalate metabolites and cluster membership were evaluated using logistic WQS regressions (Carrico et al., 2015) featuring a permutation test (Day et al., 2021; Day et al., 2022; Loftus et al., 2021; Loftus et al., 2022). For the WQS regressions, data were not split into training and validation sets for these models in order to maximize power, though this can bias standard error estimates in an anti-conservative direction (Borovicka et al., 2012). Therefore, we used a 200-iteration permutation test as implemented in the *wqspt* R package (Day et al., 2023) to obtain accurate p-values. When the WQS regression 95 % confidence intervals (CIs) did not include the null, we applied this permutation test to estimate a p-value (P^*), which is a more accurate measure of precision than the 95 % CIs (Day et al., 2022), though we provide the latter for reference. All phthalate measurements were decile-transformed for the WQS regression.

As an additional mixture analysis, we also ran quantile g-computation models for the same associations evaluated by the WQS regressions (Keil et al., 2020). Whereas the WQS regression evaluates direction-specific (i.e., positive or negative) mixture associations, quantile g-computation instead calculates an overall (i.e., positive plus negative) mixture coefficient and tests the null hypothesis that the overall coefficient is zero. As with the WQS regression, phthalate measurements were decile-transformed for the quantile g-computation models.

All analyses were performed in R (R Development Core Team, 2022) version 4.1.3 with the packages *clusterGeneration* (Qiu and Joe, 2020), *missMDA* (Josse and Husson, 2016), *fpc* (Hennig, 2020), *NbClust* (Charrad et al., 2014), *qgcomp* (Keil, 2022), and *wqspt* (Day et al., 2023).

3. Results

3.1. Participant characteristics

A total of 1700 ECHO-PATHWAYS mother–child dyads participated in age 8–9 visits. Of these, 1092 participants (735 CANDLE, 91 GAPPS, and 266 TIDES) had complete outcome data and were included in the clustering analysis. Comparisons of demographic and other covariate data for included and excluded participants are shown in Table S2. Demographic and other covariate data for participants in the complete data clustering analysis are described in Table 1. Of those 1092 participants, 856 participants also had complete phthalate and covariate data and were included in the phthalate regressions.

The Clustering Dataset refers to the 1092 ECHO PATHWAYS participants with complete outcome data who were then clustered based on that outcome data. The Regression Dataset is a subset of 856 of those participants in the Clustering Dataset who also have complete phthalate exposure and covariate data, and therefore were included in the regressions associating phthalates with outcome cluster membership. The third column shows missingness counts and percentages for each covariate within the Regression Dataset.

3.2. Outcome Summary Statistics

Outcome variables used for complete data clustering are described in Table 2. Complete data outcome descriptions by study site are provided in Table S1. 11.4 %, 5.0 %, 17.1 %, 9.9 %, and 35.2 % children had parent-reported speech problems, learning disabilities, asthma, wheeze, and rhinitis, respectively. Mean SCARED scores were just above the cutoff of 25 for clinically relevant anxiety (Behrens et al., 2019). Pearson correlation coefficients for the residuals of outcome variables regressed on study site are shown in Figure S4. These correlations were generally low, with the highest values being between CBCL subscale scores such as between externalizing and either social problems (Pearson coefficient = 0.68) or internalizing problems (0.66) scores.

3.3. Clustering results

The most common optimal number of clusters chosen was three (Table S3). Fig. 1A shows the outcome residual z-score values for the coordinates of the three residual k-means cluster centers, and Fig. 1B summarizes the z-scores and percentages for continuous and categorical outcomes, respectively, for each of the three clusters.

Fig. 1A is a bar plot showing the z-score coordinates of each variable for each cluster center. These z-scores are the centered and scaled residuals from regressing each variable on study site. These z-scores are colored on a scale of purple for low z-scores to yellow for high z-scores. Fig. 1B is a heatmap describing the distributions of the outcome data on their original scale by cluster. In the top plot, continuous means and standard deviations are presented for each variable on its original scale. The heatmap is filled in based on the z-score of those values as a way of demonstrating relative differences for each variable, which is shown in the legend labeled “Z” on the right. The bottom plot uses a different heatmap color scale to show the percentages of each binary outcome variable with lower percentages being purple and higher ones being yellow.

Cluster 1 had the highest mean FSIQ, lowest mean BMI, lowest behavioral assessment scores, and lowest prevalence of all adverse conditions than the other two clusters except speech problems, which had a lower frequency in Cluster 3 (6 % vs. 3.6 %, respectively). Cluster 2 (n = 192) had the lowest mean FSIQ and highest prevalence of every adverse outcome except asthma, wheeze, and rhinitis, though those frequencies were still higher than in Cluster 1. Cluster 3 had the highest frequencies of asthma (92 %), wheeze (54 %), and rhinitis (57 %), as well as slightly elevated mean BMI, SCARED score (Anxiety), and CDI score (Depression) z-scores compared to Cluster 1. Both Clusters 2 (28.6 %) and 3 (29.5 %) had higher obesity rates than Cluster 1 (17.0 %). The mean bootstrap Jaccard indices were 0.88, 0.83, 0.71 for Clusters 1, 2, and 3, respectively, indicating that Clusters 1 and 2 are stable or highly stable, while Cluster 3 is moderately stable (Hennig, 2008).

Differences in demographic variables by complete data cluster are summarized in Table S4. Compared to Cluster 1, children in Clusters 2 and 3 were more likely to be male (Clusters 2/3: 59 % vs. Cluster 1: 45 % male) and have mothers with higher pre-pregnancy BMI (mean: 28.8 vs. 27.1) and lower educational attainment (bachelor's degree: 40 % vs. 55 %). Maternal stress was highest in Cluster 2 (Cluster 1 mean PSS-4 score: 5.0, Cluster 2: 6.6, Cluster 3: 5.2). Cluster 3 had the lowest percentage of non-Hispanic White mothers (Cluster 1: 44 %, Cluster 2: 41 %, Cluster 3: 22 %), lowest income (Cluster 1 mean: \$66 k, Cluster 2: \$53 k, Cluster 3: \$47 k), highest percentage of mothers who smoked (Cluster 1: 7.1 %, Cluster 2: 12 %, Cluster 3: 14 %) or had asthma (Cluster 1: 11 %, Cluster 2: 21 %, Cluster 3: 32 %), and more CANDLE participants (Cluster 1: 65 %, Cluster 2: 67 %, Cluster 3: 80 %).

3.4. WQS regressions

Summary statistics and correlations for the phthalate concentrations are shown in Table S6 and Figure S5. Fig. 2 shows odds ratios, full sample 95 % CIs, and permutation-test p-values (P^*) from the WQS regressions, and Table S7 shows mixture weights. In the sex-stratified analysis in males, odds of Cluster 3 membership were significantly positively associated with a mixture of 15 prenatal phthalate metabolites (OR: 1.59 (95 % CI: 1.26, 2.03), $P^*=0.01$). This mixture had high weights for monoethyl phthalate (MEP) (weight = 0.38), monoheptyl phthalate (MHPP) (0.18), and the DEHP metabolites (summed weights = 0.14). Female-specific associations were null.

Fig. 2 shows the odds ratio coefficients and 95 % confidence intervals (CIs) for associations between either Cluster 2 or 3 and Cluster 1 as determined by logistic WQS regression. The 95 % CIs output by WQS regression when not splitting data into training and validation sets (i.e., using the full sample) are prone to anticonservative bias, and so a permutation test was done to get valid p-values (P^*). As these p-values are more conservative than the default CIs would indicate, we only obtained these p-values when the full sample 95 % CIs did not overlap the null. P^* values are shown for those coefficients with full sample 95 % CIs not overlapping the null.

The additional mixture analysis results using quantile g-computation are shown in Figure S7 and Tables S10 and S11. These results were similar to the WQS regression analysis, with the only significant overall (i.e., positive direction plus negative direction) mixture coefficient

being for the association between Cluster 3 membership and the phthalate mixture in males (OR = 1.21 (1.01, 1.44), $p = 0.04$).

3.5. Individual phthalate regressions

Fig. 3 shows the regression results when individually associating 15 phthalate metabolites and the molar sum of DEHP metabolites with the odds of membership in Clusters 2 or 3 either without (A) or with (B) sex interaction terms. Without sex interaction terms, there was a significant association between Cluster 3 membership and mono(2-ethyl-5-hydroxyhexyl) phthalate (MEHHP) (OR: 1.80 (95 % CI: 1.12, 2.91)), as well as a suggestive association with MHPP (OR: 1.47 (0.95, 2.27)). In the sex interaction models, there was a significant interaction and male-specific coefficient for the association between Cluster 3 membership and MEP (Male OR: 2.51 (1.79, 4.41), interaction $p = 0.002$). Male-specific coefficients in association with Cluster 3 membership were also significant for MHPP (OR: 1.93 (1.04, 3.62)) and MEHHP (OR: 2.35 (1.53, 4.90), though without significant sex interaction terms (interaction $p = 0.32$ and 0.50 , respectively). Coefficients are shown in Tables S7-8.

These means and 95 % confidence intervals (CIs) represent odds ratios in associations between individual phthalates and either Cluster 2 versus Cluster 1 membership or Cluster 3 versus Cluster 1 membership. In both of these cases, observations assigned to the other cluster are omitted. Sample sizes for the Cluster 2 and Cluster 3 regressions were 729 and 714, respectively. Fig. 3A shows results from models without interaction terms by sex, while Fig. 3B shows results marginal sex-specific odds ratios for the models with a phthalate by sex multiplicative interaction term.

4. Discussion

In this multi-cohort, multi-site prospective study, we identified three subpopulations of children, including two with multimorbid obesity-related, neurodevelopmental, and respiratory outcomes. We observed that children sorted into Clusters 2 and 3 had lower socioeconomic status on average. In a demonstration of the application of these clusters to epidemiologic studies, we observed predominantly male-specific individual and mixture associations between MEP, MHPP, and MEHHP and membership in Cluster 3. Our study is among the first to associate childhood multi-outcome clusters with environmental exposures. Two prior latent class analyses associated socioeconomic factors with clusters of multimorbid pediatric diagnoses, with both observing broad associations between either lower socioeconomic position levels or parental education and each derived multimorbid cluster (Carrilero et al., 2020; Schramm et al., 2022). Our analysis differs in its focus on subclinical outcome assessments and its novel identification of an environmental exposure that is associated with multi-morbidity in Cluster 3. An additional recent study took an alternative multi-outcome analytic approach of deriving the first principal component from separate factor analyses of neurodevelopmental and respiratory outcome domains along with separately averaging cardio-metabolic outcomes, averaging each domain score together to get a health score, and finally associating that score with social and environmental exposures (Amine et al., 2023). Our clustering approach differs in that it captured patterns of health outcomes within and across domains instead of evaluating separate representations of each

domain. Our approach also clusters based on the full variability of the variables in each domain, whereas the Amine et al. analysis focuses on the variability in the first principal component for each domain.

Cluster 2 was characterized by higher levels of adverse neurobehavioral outcomes along with a 60 % higher obesity rate and substantially higher percentages of adverse respiratory outcomes than the healthier Cluster 1. While the asthma rate in Cluster 2 is similar to the overall rate in this population, our analysis demonstrates that the overall rate is the average of low, medium, and high respiratory outcome rates over the subpopulations in the observed clusters. Prior research found that obese children are more likely to have behavioral problems and allergic respiratory conditions (Halfon et al., 2013), and that children with asthma and rhinitis are more likely to exhibit a variety of adverse behavioral pathologies (Garg and Silverberg, 2014), in particular anxiety and depression (Vargas, 2020), as well as cognitive issues like speech disorders (Strom and Silverberg, 2016). One key pathway underlying the observed cross-domain multimorbidity may be the activation of indoleamine-2,3-dioxygenase induced by allergic responses (van der Leek et al., 2017) or obesity-related inflammation (Wolowczuk et al., 2012). This enzyme promotes compensatory immunoregulation but also leads to cerebral serotonin depletion characteristic of depression (Haroon et al., 2012).

Cluster 3 also had a similarly elevated obesity rate to Cluster 2 with a high multi-respiratory outcome phenotype when compared to the healthier Cluster 1, as well as slightly elevated levels of anxiety and depression. A previous multi-omic clustering analysis of pediatric asthma and allergies revealed a high-immunoglobulin E (IgE), poly-sensitized sub-phenotype with high respiratory multimorbidity (Anto et al., 2017). Though our analysis did not evaluate IgE, it is possible that the high respiratory outcome pattern we observed in Cluster 3 reflects a related phenotype. Prior analyses evaluating comorbidities with pediatric asthma found high (59–78 %) comorbidity with rhinitis and more limited (8–16 %) comorbidity with anxiety, depression, and obesity (Kaplan et al., 2020). Mechanisms linking asthma with anxiety and depression may include downregulation of β 2-adrenergic and glucocorticoid receptors (Miller and Chen, 2006), differential epidermal growth factor receptor-related signaling (Park et al., 2018), and perturbations to the gut microbiome and lining (Bhatt et al., 2022), though it is uncertain whether these outcomes arise simultaneously or in some temporal order (Carroll, 2014). Socioeconomic disadvantage has been associated with comorbidity of clinical diagnoses in early childhood (Carrilero et al., 2020; Russell et al., 2020), reflecting the demographic differences we observed between clusters.

The observed male-specific association of a mixture of prenatal phthalates with Cluster 3, which was defined by markedly higher rates of age 8–9 respiratory outcomes, is supported by a prior analysis of ECHO PATHWAYS Consortium data that also observed a male-specific association between a WQS regression-derived prenatal phthalate mixture highly weighted for MEP and higher odds of ever asthma at an earlier age 4–6 time point, though with a slightly lower odds ratio (Adgent et al., 2020). Other papers evaluating sex differences in early life phthalate associations with asthma or wheeze risk found male-specific prenatal and postnatal monobenzyl phthalate and mono(2-ethylhexyl) phthalate

metabolite associations with asthma, wheeze, and IgE (Ku et al., 2015); male-specific cross-sectional associations between MEP and asthma in male children (Odebeatu et al., 2019); and male-specific associations between asthma and both prenatal monoisobutyl phthalate and monoisodecyl phthalate as well as postnatal MEP and allergy (Foong et al., 2022; Navaranjan et al., 2021). However, other studies suggested only null associations or did not evaluate association modifications by sex (Johnk et al., 2020; Wu et al., 2020). Future analyses validating our results using a similar approach in an independent sample would be useful for better understanding the sex-specific nature of phthalate associations with respiratory outcomes. There are several possible mechanisms that could underly these observations. Phthalates may induce an allergic Type 2 inflammatory response through epigenetic changes to hypersensitivity-related immune factor expression, including hypomethylation of TNF α (Wang et al., 2015) and increasing CD4⁺ T cell IL-13 expression (Kuo et al., 2013). Furthermore, estrogen enhances and testosterone blocks Type 2 inflammation (Yung et al., 2018), so the pro-estrogenic and anti-androgenic activity of various phthalates may elicit an atopic phenotype (Sathyanarayana et al., 2014). Maternal testosterone during pregnancy was associated with lower IgE in boys but not girls in prior work (Shaheen et al., 2007), consistent with our sex-specific findings. This mixture analysis also showcases some of the differences between the WQS regression and quantile g-computation approaches to estimating exposure mixture associations, with the quantile g-computation overall coefficient being a slightly attenuated estimate compared to the direction-specific coefficient from WQS regression in our results. This is a result of negative associations with some mixture components partially attenuating the overall positive association of interest in the quantile g-computation analysis.

4.1. Strengths and limitations

Strengths of this study include the large racially and ethnically diverse consortium of prospective cohorts from six U.S. cities with a varied set of assessment-based health outcomes, including robust, validated assessments of neurodevelopmental outcomes. This study also used extensive simulations to validate the clustering algorithm applied in this analysis. Additionally, it utilized advanced multiple mixture exposure regression methods with a robust set of confounders to demonstrate how the determined clusters can advance pediatric health instead of traditional single outcome epidemiologic analysis.

Limitations of this study included that the cohorts of ECHO-PATHWAYS were not designed to recruit representative study populations (LeWinn et al., 2022), which may impact the generalizability of these results. Harmonization of variables across cohorts resulted in loss of information. Notably, BMI, FSIQ, speech problems, and learning disabilities were harmonized across different ages and for FSIQ even different assessments, and so age at assessment may have impacted results. However, BMI z-scores moderately correlated across visits, and studies have demonstrated the stability of IQ (Deary, 2014), speech problems (Beitchman et al., 1994), and learning disabilities (Tomblin et al., 2003) throughout childhood. The assessment of phthalate exposure relied on a spot urine sample, which may lead to exposure misclassification bias. This is because repeated phthalate measurements during pregnancy have been shown to largely have poor consistency due to their short biological half-lives (Gaylord et al., 2022). Prior evidence suggests that late pregnancy is

a key exposure window for phthalate associations with several of the outcomes included in this analysis, but it is possible that there are other relevant exposure windows or that differences in critical windows of exposure between the different outcomes could impair our ability to detect an association using our multi-outcome clustering approach. In addition to focusing on multi-outcome methods, we demonstrated an application of mixture exposure epidemiology, which can result in model estimates impacted by complex causal structures that are not addressed by WQS regression or similar methods (Weisskopf et al., 2018). Our selection of outcome variables included more variables for neurodevelopment than the other disease domains, which could potentially lead to clusters that are unequally determined by each of our domains of interest. Finally, our approach assumes a common directionality of exposure associations with outcomes overrepresented in a given cluster, and opposing or primarily null associations with those individual outcomes could mask associations with cluster membership.

5. Conclusions

We used a multi-outcome epidemiologic approach that leverages wide-ranging outcome data collection to identify subpopulations of children, finding distinct patterns of multi-domain multimorbidity that included clustered respiratory outcomes associated with prenatal phthalate exposures in boys. Policy efforts should reduce phthalate exposures in mothers to limit respiratory disease and more broadly focus on subpopulations of children experiencing clusters of co-occurring pathological symptoms. Future research can further refine these multi-morbid phenotypes by examining patterns among similar outcomes in other multi-cohort populations and using biomarkers to determine shared mechanisms. By using these multi-outcome phenotypes in future environmental epidemiologic studies, researchers can identify public interventions with maximal impact.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank our ECHO Colleagues; the medical, nursing, and program staff; and the children and families participating in the ECHO cohorts. We also acknowledge the contribution of the following ECHO Program collaborators:

ECHO Components—Coordinating Center: Duke Clinical Research Institute, Durham, North Carolina: Smith PB, Newby LK; Data Analysis Center: Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland: Jacobson LP; Research Triangle Institute, Durham, North Carolina: Catellier DJ; Person-Reported Outcomes Core: North-western University, Evanston, Illinois: Gershon R, Cella D.

Funding/Support

This study was supported by the Environmental influences on Child Health Outcomes (ECHO) Program, Office of the Director, National Institutes of Health, under Award Numbers U2COD023375 (Coordinating Center), U24OD023382 (Data Analysis Center), and U24OD023319 with co-funding from the Office of Behavioral and Social Science Research (PRO Core). This research was also funded by the ECHO Program Opportunities and Infrastructure Fund grant U2COD023375 (Day, Szpiro), as well as the ECHO-PATHWAYS consortium (NIH grants: UG300023271/UH3OD023271 (Karr) and UH3OD023305 (Trasande)). The CANDLE study is also funded by the Urban Child Institute, NIH grant 1R01HL109977, R01HL132338, and CIHR award number MWG-146331. The TIDES study is also supported by the NIH Office of Director (UG3OD023305 and UH3OD023305), NIEHS

intramural funding ZIA10331, and NIH grants R01ES016863 and R01ES25169. Dr. Barrett was additionally funded by NIEHS grant P30ES005022.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability

Data will be made available on request.

Abbreviations:

ARI	Affective Reactivity Index
ARI	adjusted Rand index
BMI	body mass index
CANDLE	Conditions Affecting Neurocognitive Development and Learning in Early childhood
CBCL	Child Behavior Checklist
CDI	Children's Depression Inventory 2 (CDI)
DEHP	di(2-ethylhexyl) phthalate
FSIQ	full-scale intelligence quotient
GAPPS	Global Alliance to Prevent Prematurity and Stillbirth
ISAAC	International Study of Asthma and Allergies in Childhood
MEP	monoethyl phthalate
MHPP	monoheptyl phthalate
MEHHP	mono(2-ethyl-5-hydroxyhexyl) phthalate
PSS-4	Perceived Stress Scale
SCARED	Screen for Child Anxiety Related Emotional Disorders
TIDES	The Infant Development and the Environment Study
WQS	weighted quantile sum.

References

- Achenbach TM. 2001. Manual for ASEBA school-age forms & profiles. University of Vermont, Research Center for Children, Youth & Families.
- Adgent MA, Carroll KN, Hazlehurst MF, Loftus CT, Szpiro AA, Karr CJ, et al. , 2020. A combined cohort analysis of prenatal exposure to phthalate mixtures and childhood asthma. *Environ. Int* 143, 105970. [PubMed: 32763629]

- Alegria M, O'Malley IS, DiMarzio K, Zhen-Duan J, 2022. Framework for understanding and addressing racial and ethnic disparities in children's mental health. *Child. Adolesc. Psychiatr. Clin. N. Am* 31, 179–191. [PubMed: 35361358]
- Amine I, Guillien A, Philippat C, Anguita-Ruiz A, Casas M, de Castro M, et al. , 2023. Environmental exposures in early-life and general health in childhood. *Environ. Health* 22, 53. [PubMed: 37480033]
- Anto JM, Bousquet J, Akdis M, Auffray C, Keil T, Momas I, et al. , 2017. Mechanisms of the Development of Allergy (MeDALL): Introducing novel concepts in allergy phenotypes. *J. Allergy Clin. Immunol* 139, 388–399. [PubMed: 28183433]
- Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. , 1995. International Study of Asthma and Allergies in Childhood (Isaac) - Rationale and methods. *Eur. Respir. J* 8, 483–491. [PubMed: 7789502]
- Attina TM, Malits J, Naidu M, Trasande L, 2019. Racial/ethnic disparities in disease burden and costs related to exposure to endocrine-disrupting chemicals in the United States: an exploratory analysis. *J. Clin. Epidemiol* 108, 34–43. [PubMed: 30529005]
- Barrett ES, Sathyanarayana S, Janssen S, Redmon JB, Nguyen RHN, Kobrosly R, et al. , 2014. Environmental health attitudes and behaviors: findings from a large pregnancy cohort study. *Eur. J. Obstet. Gynecol. Reprod. Biol* 176, 119–125. [PubMed: 24647207]
- Behrens B, Swetlitz C, Pine DS, Pagliaccio D, 2019. The Screen for Child Anxiety Related Emotional Disorders (SCARED): Informant discrepancy, measurement invariance, and test-retest reliability. *Child. Psychiatry Hum. Dev* 50, 473–482. [PubMed: 30460424]
- Beitchman JH, Brownlie EB, Inglis A, Wild J, Mathews R, Schachter D, et al. , 1994. Seven-year follow-up of speech/language-impaired and control children: speech/language stability and outcome. *J. Am. Acad. Child. Adolesc. Psychiatry* 33, 1322–1330. [PubMed: 7995800]
- Bhatt S, Pai KSR, Patil C, Manjula S, Mohana Lakshmi S. 2022. Role of Brain–Gut–Microbiome Axis in Depression Comorbid with Asthma. In: *Microbiome in Inflammatory Lung Diseases*:Springer, 135–151.
- Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, et al. , 1997. The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. *J. Am. Acad. Child. Adolesc. Psychiatry* 36, 545–553. [PubMed: 9100430]
- Borovicka T, Jirina M Jr, Kordik P, Jirina M, 2012. Selecting representative data sets. *Adv. Data Mining Knowledge Discov. Appl* 12, 43–70.
- Buckley JP, Doherty BT, Keil AP, Engel SM, 2017. Statistical approaches for estimating sex-specific effects in endocrine disruptors research. *Environ. Health Persp* 125.
- Carrico C, Gennings C, Wheeler DC, Factor-Litvak P, 2015. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. *J. Agric. Biol. Environ. Stat* 20, 100–120. [PubMed: 30505142]
- Carrilero N, Dalmau-Bueno A, Garcia-Altes A. 2020. Comorbidity patterns and socioeconomic inequalities in children under 15 with medical complexity: a population-based study. *Bmc Pediatr* 20.
- Carroll KN, 2014. Can the blues make it harder to breathe? *Am. J. Resp. Crit. Care* 189, 1013–1014.
- Charrad M, Ghazzali N, Boiteau V, Niknafs A, 2014. NbClust: an R package for determining the relevant number of clusters in a data set. *J. Stat. Softw* 61, 1–36.
- Cohen S, Kamarck T, Mermelstein R. 1983. A global measure of perceived stress. *J Health Soc Behav*:385–396. [PubMed: 6668417]
- Davis CM, Apter AJ, Casillas A, Foggs MB, Louisias M, Morris EC, et al. , 2021. Health disparities in allergic and immunologic conditions in racial and ethnic underserved populations: A Work Group Report of the AAAAI Committee on the Underserved. *J. Allergy Clin. Immunol* 147, 1579–1593. [PubMed: 33713767]
- Day DB, Collett BR, Barrett ES, Bush NR, Swan SH, Nguyen RHN, et al. , 2021. Phthalate mixtures in pregnancy, autistic traits, and adverse childhood behavioral outcomes. *Environ. Int* 147.
- Day DB, Sathyanarayana S, LeWinn KZ, Karr CJ, Mason WA, Szpiro AA, 2022. A permutation test-based approach to strengthening inference on the effects of environmental mixtures: Comparison between single -index analytic methods. *Environ. Health Perspect* 130.

- Day DB, Peng J, Szpiro AA. 2023. wqspt: Permutation test for weighted quantile sum regression. R package version 1.0.1.
- Deary IJ, 2014. The stability of intelligence from childhood to old age. *Curr. Dir. Psychol. Sci* 23, 239–245.
- Eales J, Bethel A, Galloway T, Hopkinson P, Morrissey K, Short RE, et al. , 2022. Human health impacts of exposure to phthalate plasticizers: An overview of reviews. *Environ. Int* 158.
- Foong RE, Franklin P, Sanna F, Hall GL, Sly PD, Thorstensen EB, et al. , 2022. Longitudinal effects of prenatal exposure to plastic-derived chemicals and their metabolites on asthma and lung function from childhood into adulthood. *Respirology*.
- Garg N, Silverberg JI, 2014. Association between childhood allergic disease, psychological comorbidity, and injury requiring medical attention. *Ann. Allerg. Asthma Im* 112, 525–532.
- Gaylord A, Kannan K, Lakuleswaran M, Zhu HK, Ghassabian A, Jacobson MH, et al. , 2022. Variability and correlations of synthetic chemicals in urine from a New York City-based cohort of pregnant women. *Environ. Pollut* 309.
- Guo Y, Weck J, Sundaram R, Goldstone AE, Louis GB, Kannan K, 2014. Urinary concentrations of phthalates in couples planning pregnancy and its association with 8-hydroxy-2'-deoxyguanosine, a biomarker of oxidative stress: Longitudinal investigation of fertility and the environment study. *Environ. Sci. Technol* 48, 9804–9811. [PubMed: 25068827]
- Halfon N, Larson K, Slusser W, 2013. Associations between obesity and comorbid mental health, developmental, and physical health conditions in a nationally representative sample of US children aged 10 to 17. *Acad. Pediatr* 13, 6–13. [PubMed: 23200634]
- Haroon E, Raison CL, Miller AH, 2012. Psychoneuroimmunology meets neuropsychopharmacology: Translational implications of the impact of inflammation on behavior. *Neuropsychopharmacol.* 37, 137–162.
- Hartigan JA, Wong MA, 1979. Algorithm AS 136: A k-means clustering algorithm. *J. Roy. Stat. Soc.: Ser. C (Appl. Stat.)* 28, 100–108.
- Hauser R, Calafat AM, 2005. Phthalates and human health. *Occup. Environ. Med* 62.
- Hennig C, 2007. Cluster-wise assessment of cluster stability. *Comput. Stat. Data Anal* 52, 258–271.
- Hennig C, 2008. Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. *J. Multivariate Anal* 99, 1154–1176.
- Hennig C, 2020. fpc: Flexible Procedures for Clustering. R package version 2.2-9.
- Hornung RW, Reed LD, 1990. Estimation of average concentration in the presence of nondetectable values. *Appl. Occup. Environ. Hyg* 5, 46–51.
- Hubert L, Arabie P, 1985. Comparing partitions. *J. Classif* 2, 193–218.
- Isong IA, Rao SR, Bind MA, Avendano M, Kawachi I, Richmond TK, 2018. Racial and ethnic disparities in early childhood obesity. *Pediatrics* 141.
- Johnk C, Host A, Husby S, Schoeters G, Timmermann CAG, Kyhl HB, et al. , 2020. Maternal phthalate exposure and asthma, rhinitis and eczema in 552 children aged 5 years; a prospective cohort study. *Environ. Health-Glob* 19.
- Josse J, Husson F, 2016. missMDA: A package for handling missing values in multivariate data analysis. *J. Stat. Softw* 70.
- Kannan K, Stathis A, Mazzella MJ, Andra SS, Barr DB, Hecht SS, et al. , 2021. Quality assurance and harmonization for targeted biomonitoring measurements of environmental organic chemicals across the Children's Health Exposure Analysis Resource laboratory network. *Int. J. Hyg. Environ. Health* 234.
- Kaplan A, Szeffler SJ, Halpin DMG, 2020. Impact of comorbid conditions on asthmatic adults and children. *NPJ Prim. Care Respir. Med* 30, 36. [PubMed: 32820164]
- Keil AP, 2022. qgcomp: Quantile G-computation. R package version 2.8.6.
- Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ, 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. *Environ. Health Perspect* 128, 47004. [PubMed: 32255670]
- Kim SH, Park MJ, 2014. Phthalate exposure and childhood obesity. *Ann. Pediatr. Endocrinol. Metab* 19, 69–75. [PubMed: 25077088]

- Knapp EA, Kress AM, Parker CB, Page GP, McArthur K, Gachigi KK, et al. , 2023. The Environmental Influences on Child Health Outcomes (ECHO)-wide cohort. *Am. J. Epidemiol* 192, 1249–1263. [PubMed: 36963379]
- Kovacs M. 1992. Children's Depression Inventory (CDI): technical manual update:Multi-Health Systems.
- Ku HY, Su PH, Wen HJ, Sun HL, Wang CJ, Chen HY, et al. , 2015. Prenatal and postnatal exposure to phthalate esters and asthma: a 9-year follow-up study of a taiwanese birth cohort. *PLoS One* 10, e0123309. [PubMed: 25875379]
- Kuczumski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, et al. , 2002. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat.* 11, 1–190.
- Kuo CH, Hsieh CC, Kuo HF, Huang MY, Yang SN, Chen LC, et al. , 2013. Phthalates suppress type I interferon in human plasmacytoid dendritic cells via epigenetic regulation. *Allergy* 68, 870–879. [PubMed: 23738920]
- Levine L, 1945. Evaluation of urinary lead determinations. I. The significance of the specific gravity. *J. Ind. Hyg. Toxicol* 27, 217–223.
- LeWinn KZ, Karr CJ, Hazlehurst M, Carroll K, Loftus C, Nguyen R, et al. , 2022. Cohort profile: the ECHO prenatal and early childhood pathways to health consortium (ECHO-PATHWAYS). *BMJ Open* 12.
- Li MC, Chen CH, Guo YL, 2017. Phthalate esters and childhood asthma: A systematic review and congener-specific meta-analysis. *Environ. Pollut.* 229, 655–660. [PubMed: 28692937]
- Loftus CT, Bush NR, Day DB, Ni Y, Tylavsky FA, Karr CJ, et al. , 2021. Exposure to prenatal phthalate mixtures and neurodevelopment in the Conditions Affecting Neurocognitive Development and Learning in Early childhood (CANDLE) study. *Environ. Int* 150.
- Loftus CT, Szpiro AA, Workman T, Wallace ER, Hazlehurst MF, Day DB, et al. , 2022. Maternal exposure to urinary polycyclic aromatic hydrocarbons (PAH) in pregnancy and childhood asthma in a pooled multi-cohort study. *Environ. Int* 170.
- Miller GE, Chen E, 2006. Life stress and diminished expression of genes encoding glucocorticoid receptor and beta(2)-adrenergic receptor in children with asthma. *P. Natl. Acad. Sci. USA* 103, 5496–5501.
- Navaranjan G, Diamond ML, Harris SA, Jantunen LM, Bernstein S, Scott JA, et al. 2021. Early life exposure to phthalates and the development of childhood asthma among Canadian children. *Environ Res* 197.
- Neier K, Marchlewicz EH, Dolinoy DC, Padmanabhan V. 2015. Assessing Human Health Risk to Endocrine Disrupting Chemicals: a Focus on Prenatal Exposures and Oxidative Stress. *Endocr Disruptors (Austin)* 3.
- Ni Y, Loftus CT, Szpiro AA, Young MT, Hazlehurst MF, Murphy LE, et al. , 2022. Associations of pre- and postnatal air pollution exposures with child behavioral problems and cognitive performance: A U.S. multi-cohort study. *Environ. Health Perspect.* 130 (67008).
- Odebeatu CC, Taylor T, Fleming LE, Osborne NJ, 2019. Phthalates and asthma in children and adults: US NHANES 2007–2012. *Environ. Sci. Pollut. R* 26, 28256–28269.
- Park HW, Song WJ, Cho SH, McGeachiel MJ, Martinez F, Mauger D, et al. 2018. Assessment of genetic factor and depression interactions for asthma symptom severity in cohorts of childhood and elderly asthmatics. *Exp Mol Med* 50.
- Perou R, Bitsko RH, Blumberg SJ, Pastor P, Ghandour RM, Gfroerer JC, et al. , 2013. Mental health surveillance among children - United States, 2005–2011. *MMWR Morb. Mortal. Wkly. Rep* 62, 1–35. [PubMed: 23302815]
- Perrin JM, Bloom SR, Gortmaker SL, 2007. The increase of childhood chronic conditions in the United States. *JAMA* 297, 2755–2759. [PubMed: 17595277]
- Qiu W, Joe H, 2006. Generation of random clusters with specified degree of separation. *J. Classif* 23, 315–334.
- Qiu W, Joe H, 2020. clusterGeneration: Random cluster generation (with Specified Degree of Separation). R package version 1.3.7.
- R Development Core Team. 2022. R: A Language and Environment for Statistical Computing. Version 4.1.3. Vienna, Austria.:R Foundation for Statistical Computing.

- Radke EG, Braun JM, Nachman RM, Cooper GS, 2020. Phthalate exposure and neurodevelopment: A systematic review and meta-analysis of human epidemiological evidence. *Environ. Int* 137.
- Rocha BA, Asimakopoulos AG, Barbosa F, Kannan K, 2017. Urinary concentrations of 25 phthalate metabolites in Brazilian children and their association with oxidative DNA damage. *Sci. Total Environ* 586, 152–162. [PubMed: 28174045]
- Roid GH, Barram RA. 2004. *Essentials of Stanford-Binet intelligence scales (SB5) assessment*: John Wiley & Sons.
- Russell J, Grant CC, Morton SMB, 2020. Multimorbidity in early childhood and socioeconomic disadvantage: Findings from a large New Zealand child cohort. *Acad. Pediatr* 20, 619–627. [PubMed: 31574311]
- Sathyanarayana S, Barrett E, Butts S, Wang C, Swan SH, 2014. Phthalate exposure and reproductive hormone concentrations in pregnancy. *Reproduction* 147, 401–409. [PubMed: 24196015]
- Schick SF, Blount BC, Jacob PR, Saliba NA, Bernert JT, El Hellani A, et al. , 2017. Biomarkers of exposure to new and emerging tobacco delivery products. *Am. J. Physiol. Lung Cell. Mol. Physiol* 313, L425–L452. [PubMed: 28522563]
- Schramm S, Moller SP, Tolstrup JS, Laursen B, 2022. Effects of individual and parental educational levels on multimorbidity classes: a register-based longitudinal study in a Danish population. *BMJ Open* 12.
- Shaheen SO, Hines M, Newson RB, Wheeler M, Herrick DRM, Strachan DP, et al. , 2007. Maternal testosterone in pregnancy and atopic outcomes in childhood. *Allergy* 62, 25–32. [PubMed: 17156338]
- Silva MJ, Reidy JA, Herbert AR, Preau JL Jr., Needham LL, Calafat AM, 2004. Detection of phthalate metabolites in human amniotic fluid. *Bull. Environ. Contam. Toxicol* 72, 1226–1231. [PubMed: 15362453]
- Smyth RL, Openshaw PJM, 2006. Bronchiolitis. *Lancet* 368, 312–322. [PubMed: 16860701]
- Sontag-Padilla L, Burns RM, Shih RA, Griffin BA, Martin LT, Chandra A, et al. 2015. *The Urban Child Institute CANDLE Study: Methodological Overview and Baseline Sample Description*. Santa Monica, CA: RAND Corporation.
- Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, et al. , 2012. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *J. Child. Psychol. Psychiatry* 53, 1109–1117. [PubMed: 22574736]
- Strom MA, Silverberg JI, 2016. Asthma, hay fever, and food allergy are associated with caregiver-reported speech disorders in US children. *Pediatr. Allerg. Imm.-Uk* 27, 604–611.
- Swan SH, Sathyanarayana S, Barrett ES, Janssen S, Liu F, Nguyen RHN, et al. , 2015. First trimester phthalate exposure and anogenital distance in newborns. *Hum. Reprod* 30, 963–972. [PubMed: 25697839]
- Tomblin JB, Zhang X, Buckwalter P, O'Brien M, 2003. The stability of primary language disorder: four years after kindergarten diagnosis. *J. Speech Lang. Hear. Res* 46, 1283–1296. [PubMed: 14700355]
- van Buuren S., 2018. *AGD: Analysis of growth data*. R package version 0.39.
- Van Cleave J, Gortmaker SL, Perrin JM, 2010. Dynamics of obesity and chronic health conditions among children and youth. *JAMA* 303, 623–630. [PubMed: 20159870]
- van der Leek AP, Yanishevsky Y, Kozyrskyj AL. 2017. The Kynurenine Pathway As a Novel Link between Allergy and the Gut Microbiome. *Front Immunol* 8.
- VanderWeele TJ, 2017. Outcome-wide epidemiology. *Epidemiology* 28, 399–402. [PubMed: 28166102]
- Vargas PA, 2020. Spreading the word: Comorbidity of asthma and depression is not just the product of a vulnerable personality. *J. Aller. Cl. Imm.-Pract* 8, 208–209.
- Wallace ER, Buth E, Szpiro AA, Ni Y, Loftus CT, Masterson E, et al. , 2023. Prenatal exposure to polycyclic aromatic hydrocarbons is not associated with behavior problems in preschool and early school-aged children: A prospective multi-cohort study. *Environ. Res* 216, 114759. [PubMed: 36370819]
- Wang IJ, Karmaus WJJ, Chen SL, Holloway JW, Ewart S, 2015. Effects of phthalate exposure on asthma may be mediated through alterations in DNA methylation. *Clin. Epigenet* 7.

- Wechsler D. 2012. Wechsler preschool and primary scale of intelligence—fourth edition. The Psychological Corporation San Antonio, TX.
- Wechsler D. 2014. WISC-V: Technical and interpretive manual:NCS Pearson, Incorporated.
- Weisskopf MG, Seals RM, Webster TF, 2018. Bias amplification in epidemiologic analysis of exposure to mixtures. *Environ. Health Persp* 126.
- Wolowczuk I, Hennart B, Leloire A, Bessede A, Soichot M, Taront S, et al. , 2012. Tryptophan metabolism activation by indoleamine 2,3-dioxygenase in adipose tissue of obese women: an attempt to maintain immune homeostasis and vascular tone. *Am. J. Physiol.-Reg. I* 303, R135–R143.
- Wu WX, Wu CY, Ji CW, Diao FQ, Peng JL, Luo D, et al. , 2020. Association between phthalate exposure and asthma risk: A meta-analysis of observational studies. *Int. J. Hyg. Envir. Heal* 228.
- Yung JA, Fuseini H, Newcomb DC, 2018. Hormones, sex, and asthma. *Ann. Allerg. Asthma Im* 120, 488–494.
- Zota AR, Calafat AM, Woodruff TJ, 2014. Temporal trends in phthalate exposures: Findings from the National Health and Nutrition Examination Survey, 2001–2010. *Environ. Health Perspect* 122, 235–241. [PubMed: 24425099]

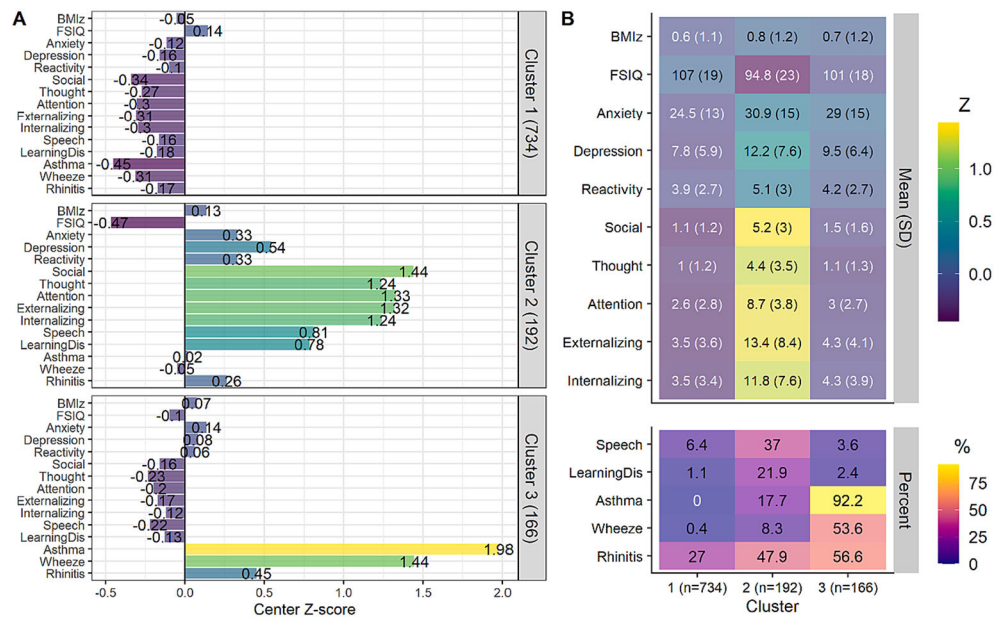


Fig. 1. Outcome Variable Cluster Center Coordinates (A) and Summaries by Cluster (B).

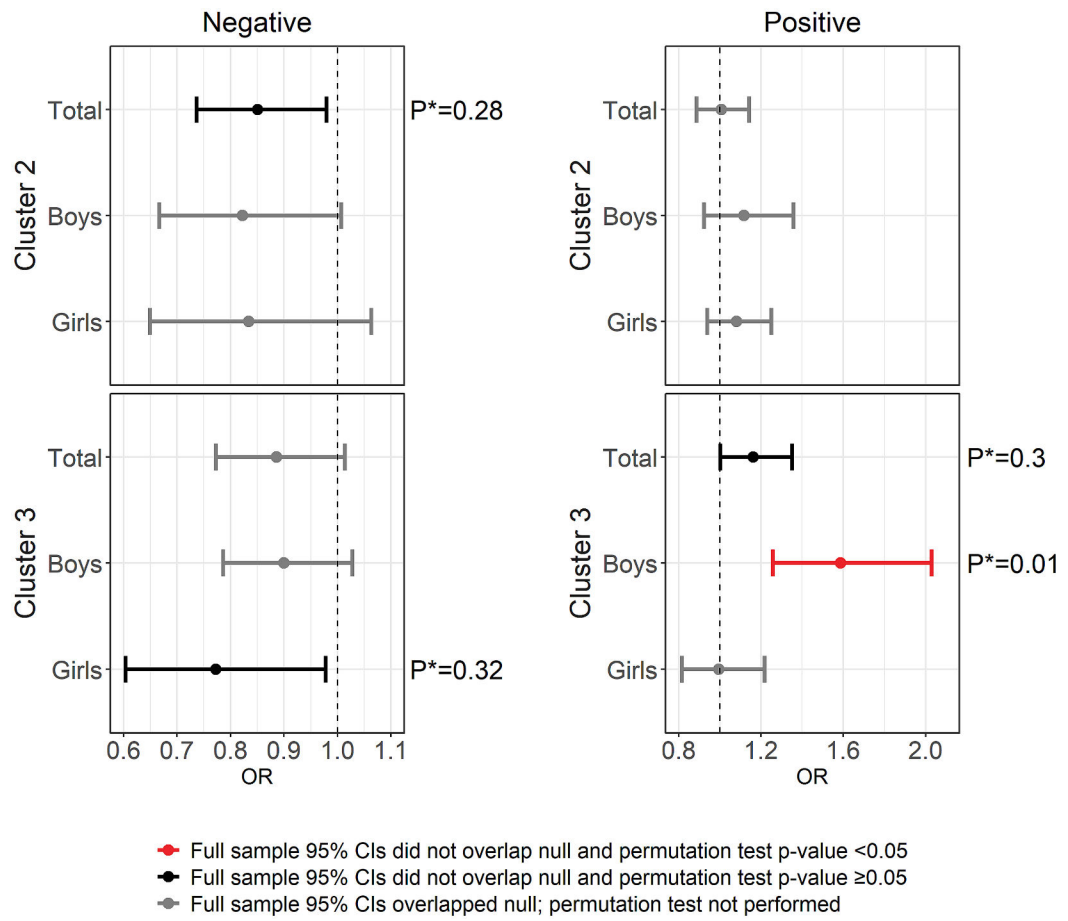


Fig. 2. Odds Ratios, Full Sample 95% Confidence Intervals, and Permutation Test-Based P-Values (P*) from Logistic Weighted Quantile Sum Regressions Associating Prenatal Phthalates and Outcome Cluster Membership.

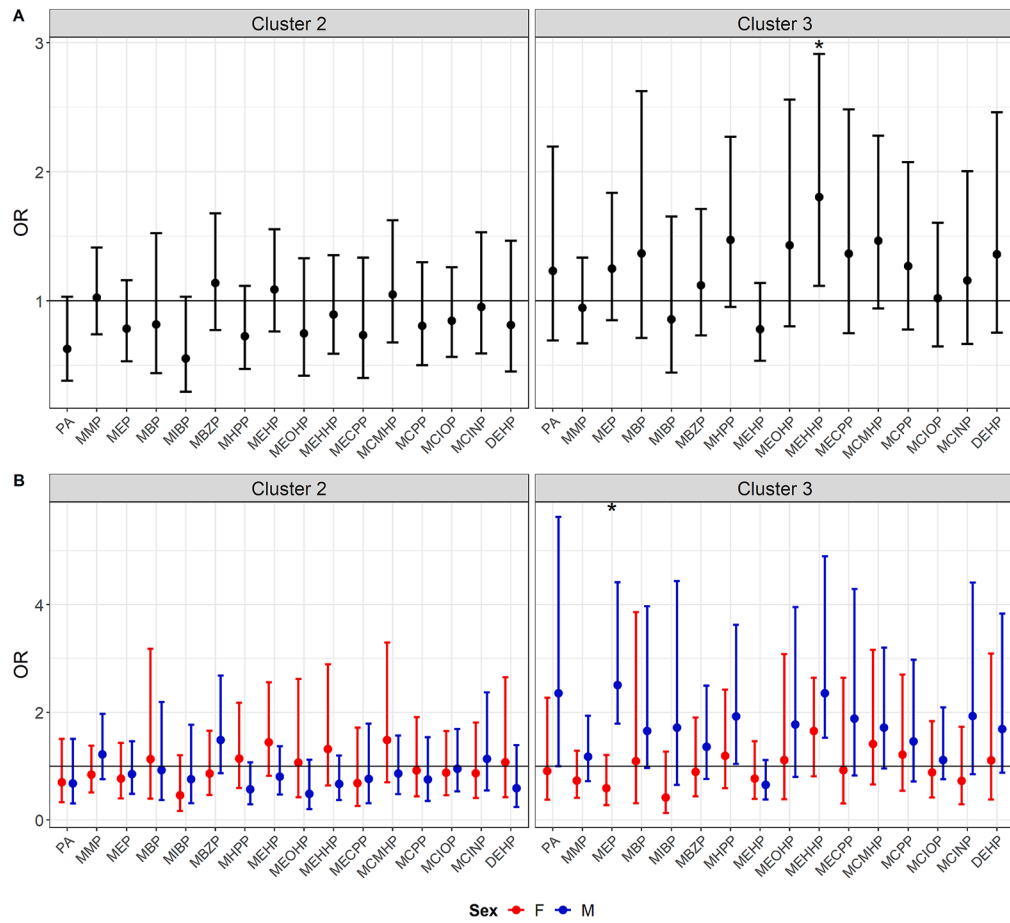


Fig. 3. Forest Plots of Mean and 95% CIs for Odds Ratios from Logistic Regressions Associating Individual Phthalates and either Cluster 2 or Cluster 3 Membership as Compared to Cluster 1 in either without (A) or with (B) Effect Modification by Sex.

Table 1

Demographic Variable and Other Covariate Summary Statistics for Clustering Dataset or the Regression Dataset.

Variable	Clustering Dataset (N = 1092)		Regression Dataset (N = 856)
	Mean (SD) or Count (%)	Count (%) Missing	Mean (SD) or Count (%)
Maternal Age	28.51 (6.04)	0 (0 %)	29.01 (5.77)
Adjusted Household Income (\$)	61,050 (51256)	36 (3.3 %)	63,881 (50924)
Pre-pregnancy BMI	27.66 (7.66)	10 (0.9 %)	27.63 (7.50)
PSS-4 Total Score	5.31 (2.86)	10 (0.9 %)	5.25 (2.84)
Specific Gravity	1.02 (0.01)	84 (7.7 %)	1.02 (0.01)
Gestational Age at Urine Collection (w)	32.02 (2.15)	88 (8.1 %)	32.07 (2.14)
Study Site: CANDLE Memphis	735 (67.31 %)	0 (0 %)	573 (66.94 %)
TIDES San Francisco	63 (5.77 %)		56 (6.54 %)
TIDES Minneapolis	98 (8.97 %)		86 (10.05 %)
TIDES Rochester	70 (6.41 %)		51 (5.96 %)
TIDES Seattle	35 (3.21 %)		26 (3.04 %)
GAPPS Seattle	59 (5.40 %)		39 (4.56 %)
GAPPS Yakima	32 (2.93 %)		25 (2.92 %)
Child Sex: Girls	550 (50.37 %)	0 (0 %)	434 (50.70 %)
Parity: Primiparous	457 (41.85 %)	7 (0.6 %)	363 (42.41 %)
Maternal Ever Asthma Diagnosis	151 (13.83 %)	115 (10.5 %)	136 (15.89 %)
Prenatal Smoking	99 (9.07 %)	0 (0 %)	73 (8.53 %)
Ethnicity: Hispanic	53 (4.85 %)	3 (0.3 %)	41 (4.79 %)
Race: White	458 (41.94 %)	4 (0.4 %)	393 (45.91 %)
Black	521 (47.71 %)		374 (43.69 %)
Asian	18 (1.65 %)		16 (1.87 %)
Pacific Islander	0 (0 %)		0 (0 %)
American Indian	< 5 (<0.46 %)		< 5 (<0.58 %)
Multiple	72 (6.59 %)		58 (6.78 %)
Other	15 (1.37 %)		11 (1.29 %)
Education: <High School	39 (3.57 %)	4 (0.4 %)	23 (2.69 %)
Grad High School	148 (13.55 %)		100 (11.68 %)
Tech School	62 (5.68 %)		42 (4.91 %)
Some College	290 (26.56 %)		216 (25.23 %)
Grad College	271 (24.82 %)		233 (27.22 %)
Masters	207 (18.96 %)		179 (20.91 %)
Doctorate	71 (6.50 %)		63 (7.36 %)
Household Count: 2–3	229 (20.97 %)	9 (0.8 %)	191 (22.31 %)
4	423 (38.74 %)		344 (40.19 %)
5	251 (22.99 %)		200 (23.36 %)
6	180 (16.48 %)		121 (14.14 %)

Variable	Clustering Dataset (N = 1092)		Regression Dataset (N = 856)
	Mean (SD) or Count (%)	Count (%) Missing	Mean (SD) or Count (%)
Urine Analysis Batch: 1	794 (72.71 %)	83 (7.6 %)	666 (77.80 %)
2	59 (5.40 %)		54 (6.31 %)
3	156 (14.29 %)		136 (15.89 %)
Birth Year: 2007	43 (3.94 %)	0 (0 %)	10 (1.17 %)
2008	139 (12.73 %)		92 (10.75 %)
2009	186 (17.03 %)		153 (17.87 %)
2010	195 (17.86 %)		169 (19.74 %)
2011	277 (25.37 %)		230 (26.87 %)
2012	185 (16.94 %)		147 (17.17 %)
2013	65 (5.95 %)		53 (6.19 %)
2014	< 5 (<0.46 %)		<5 (<0.46 %)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Summary Statistics for the 15 Outcome Variables included in the Clustering Analysis among 1092 Participants with Complete Outcome Data.

Variable (Abbreviation)	Mean (SD) or Count (%)	Median (Range)
BMI z-score (BMIz)	0.62 (1.11)	0.58 (-3.62–3.29)
Full scale IQ (FSIQ)	103.75 (19.91)	106 (18–157)
SCARED Total Score (Anxiety)	26.33 (14.29)	25 (0–80)
CDI Total Score (Depression)	8.84 (6.49)	7 (0–41)
ARI Total Score (Reactivity)	4.15 (2.77)	4 (0–12)
CBCL Social Problems Score (Social)	1.86 (2.32)	1 (0–16)
CBCL Thought Problems Score (Thought)	1.62 (2.26)	1 (0–20)
CBCL Attention Problems Score (Attention)	3.76 (3.74)	3 (0–19)
CBCL Externalizing Score (Externalizing)	5.33 (6.13)	3 (0–47)
CBCL Internalizing Score (Internalizing)	5.06 (5.47)	3.5 (0–42)
Ever Speech Problems (Speech)	124 (11.36 %)	
Ever Learning Disability (LearningDis)	54 (4.95 %)	
Ever Asthma (Asthma)	187 (17.12 %)	
Past-Year Wheeze (Wheeze)	108 (9.89 %)	
Ever Allergic Rhinitis (Rhinitis)	384 (35.16 %)	

This table summarizes the 15 outcome variables included in the complete outcome clustering analysis. Either means and standard deviations for continuous variables or counts and percentages for categorical variables are shown in the second column. The third column shows medians and ranges for only the continuous variables.