Do small effects matter more in vulnerable populations? An investigation using Environmental Influences on Child Health Outcomes (ECHO) cohorts

Additional File

eTable1 List of cohorts included

a) Cohorts providing data for birthweight and maternal pre-pregnancy BMI
43rd Multicenter Airway Research Collaboration (MARC-43)
Archive for Research in Child Health (ARCH)
Asthma Coalition on Community, Environment & Social Stress (ACCESS)
Atlanta ECHO Cohort of Emory University
BAMBAM
Boricua Youth Study (BYS)
Chemicals in our Bodies (CIOB)
Childhood Allergy and the Neonatal Environment (CANOE)
Childhood Origins of Asthma Study (COAST)
Cincinnati Childhood Allergy & Air Pollution Study (CCAAPS)
Columbia Center for Children's Environmental Health (CCCEH) - Mothers and Newborns (M and N)
Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE)
Early Autism Risk Longitudinal Investigation (EARLI)
Early Growth and Development Study Pediatric Cohort
ECHO in Puerto Rico (PROTECT)
Fair Start
Family Life Project (FLP)
First 1000 Days
Healthy Start
Illinois Kids Development Study (IKIDS)
Infant Brain Imaging Study (IBIS)
Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure (INSPIRE)
In-Utero Smoke, Vitamin C, and Newborn Lung Function
Kaiser Permanente Research Bank (KPRB)
Kennedy Krieger-Baby Siblings Research Consortium (Kennedy Krieger-BSRC)
Magee
Maternal and Development Risks from Environmental and Social Stressors (MADRES)
Michigan Archive for Research in Child Health (MARCH)
Microbes, Allergy, Asthma & Pets Study (MAAP)
MINNIE
New Hampshire Birth Cohort Study (NHBCS)
Pittsburgh Girls Study (PGS)
Pregnancy Environment and Lifestyle Study (PETALS)
PRogramming of Intergenerational Stress Mechanisms (PRISM)
Project Viva
Revisiting CHildhood Autism Risks from Genes and the Environment Study (ReCHARGE)
Rochester
Safe Passage Study (PASS)
Sibling Cohort
The Global Alliance to Prevent Prematurity and Stillbirth (GAPPS)
The Infant Development and the Environment Study (TIDES)

The NYU Children's Health and Environment Study (NYU CHES)
University of California Davis- Baby Siblings Research Consortium (UCDavis-BSRC)
University of California- Markers of Autism Risk in Babies (MARBLES)
Urban Environment and Childhood Asthma (URECA)
Utah Children’s Project
Vitamin C to Decrease Effects of Smoking in Pregnancy on Infant Lung Function (VCSIP)
Vitamin D Antenatal Asthma Reduction Trial (VDAART)
Wayne County Health Environment Allergy and Asthma (WHEALS)
Wisconsin Infant Study Cohort (WISC)

eMethods1: Study Sample

This study used data collected as part of the Environmental influences on Child Health Outcomes (ECHO) Program (cycle 1). ECHO is a consortium of pediatric cohort studies funded by NIH and established in 2016 to investigate the impacts of a wide range of environmental exposures on children's health http://www.echochildren.org/.

In cycle 1, 2016-2023, the ECHO program included 69 cohorts in 31 consortia representing the diversity in sociodemographic features and exposures among children born in the United States. The cycle 1 program includes participant data for over 43000 pregnancies and children who have consented to the ECHO-wide protocol. eTable 1 lists the participating cohorts.

The study protocol was approved by the single ECHO institutional review board, WCG IRB. Written informed consent or parent's/guardian's permission was obtained along with child assent as appropriate, for the ECHO Cohort Data and Biospecimen Collection Protocol participation and for participation in specific study sites, WCG IRB tracking number for the ECHO-wide Protocol: 20181210. Participant data obtained from each cohort were pooled and harmonized by the central ECHO Data Analysis Center to facilitate pan-cohort "ECHO-wide" analyses as conducted in the present study.

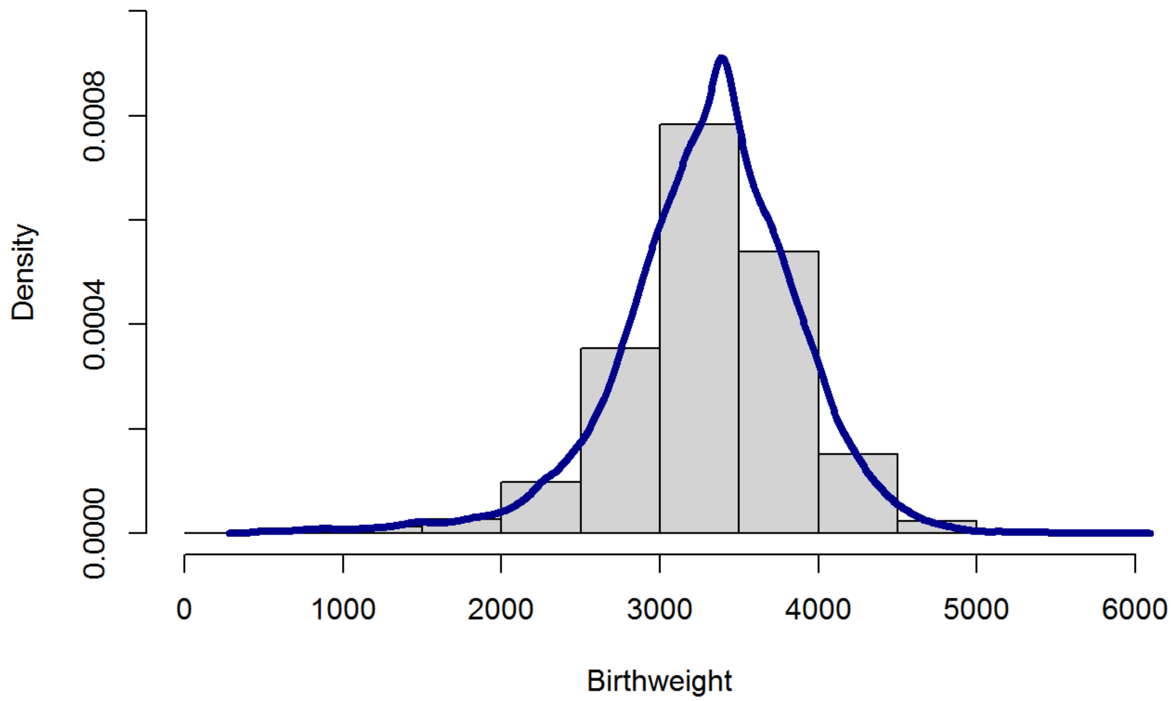
eMethods2: How the Distributional Approach works

In its simplest form that assumes the data follow a Normal distribution with equal variance in 2 groups being compared (i.e. as for a standard t test), the distributional approach works as follows: The proportion of the distribution beyond a given cut-point is calculated using the usual Normal distribution formula: i.e. proportion, P is given by Normal density (cut-point-mean)/SD. For example if (cut-point-mean)/SD = -1.96, then P =0.025. This is the way that references ranges, and in general centiles, are calculated, and a similar process is used here to calculate each proportion and the difference in proportions given the means and SDs in the 2 groups.

Some complexity comes into play in calculating the standard error of the difference of proportions as a function of the means, and this uses the delta method based on a Taylor expansion. Mathematical details of the derivation are given in our first publication[1] . The distributional method has been extended to accommodate i) unequal variances[2] , ii) skewed distributions[3] and iii) multiple regression/mixed models[4] . The software in Stata and R accommodate these extensions. Since the distributional approach uses the same regression model to estimate both the difference in means and the difference in proportions, they are testing the same hypothesis and so it is not expected that the 95% CIs will lead to different conclusions unless the underlying asymptotic assumptions break down when bootstrapped confidence intervals can be used within both the Stata[5] and R packages[6]. In addition, the method and hence the software allows the estimated difference in proportions to be reported as relative risks and odds ratios if required.

In this paper, the distributional approach has only been used to calculate the proportions beyond either a lower or upper cut-point, as defines, high-risk. The standard error and confidence intervals are not needed here but these are used when using the distributional method to calculate dual effect estimates, the difference in means and difference in proportions at high risk.

eFigure1 Distribution of birthweight in 28496 infants from the ECHO cohort consortium



eTable 2 Modeled proportions with low birthweight (<2500g) associated with, a shift in means of 167g overall and by social determinants in the ECHO consortium (N=28,496)

	% LBW unexposed	**% LBW exposed**	**Difference in percent-age points (exposed-unexposed)**	**95% Confidence Interval**
All participants	7.8	13.0	5.2	4.9, 5.5
Maternal education				
Less than high school	12.2	18.8	6.7	4.6, 8.7
High school	9.8	15.8	6.0	4.7, 7.2
College, no degree	9.9	15.7	5.8	4.9, 6.6
Bachelor's degree	7.5	12.4	4.9	4.3, 5.6
Master's degree or higher	6.5	11.2	4.7	4.1, 5.3
Missing/unknown	6.5	11.5	4.9	4.5, 5.4
Health Insurance (HI)				
No HI/public HI	9.2	15.0	5.8	5.1, 6.4
Employer/market/private	5.9	10.4	4.5	4.1, 4.9
More than one	9.5	15.2	5.7	4.7, 6.8
Missing/unknown	8.2	13.6	5.4	4.9, 6.0
Race				
White	5.9	10.3	4.4	4.1, 4.7
Black	13.5	20.9	7.4	6.3, 8.4
Asian	10.6	17.6	7.1	5.1, 9.0
Native Hawaiian/Pacific Islands	7.9	13.3	5.4	0.4, 10.5
American Indian/Alaskan Native	7.7	12.5	4.9	2.8, 6.9
Multiple Races	9.4	15.1	5.7	4.7, 6.7
Other races	9.4	15.3	5.9	4.3, 7.5
Unknown/missing	6.6	11.6	5.0	3.9, 6.1
Ethnicity				
Hispanic	8.2	13.7	5.5	4.9, 6.1
Non-Hispanic	7.6	12.7	5.1	4.8, 5.5
Unknown/missing	8.5	14.3	5.9	4.0, 7.7

eTable 3 Modeled proportions with low birthweight (<2500g) associated with, a shift in means of 250g overall and by social determinants in the ECHO consortium (N=28,496)

	% LBW unexposed	**% LBW exposed**	**Difference in percent-age points (exposed-unexposed)**	**95% Confidence Interval**
All participants	7.8	13.0	5.2	4.9, 5.5
Maternal education				
Less than high school	12.2	22.8	10.7	8.5, 12.9
High school	9.8	19.5	9.7	8.3, 11.0
College, no degree	9.9	19.2	9.3	8.4, 10.2
Bachelor's degree	7.5	15.6	8.1	7.4, 8.8
Master's degree or higher	6.5	14.3	7.8	7.1, 8.5
Missing/unknown	6.5	14.8	8.2	7.7, 8.7
Health Insurance (HI)				
No HI/public HI	9.2	18.6	9.35	8.6, 10.1
Employer/market/private	5.9	13.4	7.45	7.0, 7.9
More than one	9.5	18.8	9.26	8.1, 10.4
Missing/unknown	8.2	17.1	8.85	8.2, 9.5
Race				
White	5.9	13.2	7.3	6.9, 7.7
Black	13.5	25.3	11.8	10.6, 12.9
Asian	10.6	22.1	11.5	9.3, 13.7
Native Hawaiian/Pacific Islands	7.9	16.8	8.9	3.2, 14.5
American Indian/Alaskan Native	7.7	15.6	7.9	5.6, 10.2
Multiple Races	9.4	18.6	9.3	8.1, 10.4
Other races	9.4	19.0	9.6	7.9, 11.4
Unknown/missing	6.6	14.9	8.3	7.1, 9.5
Ethnicity				
Hispanic	8.2	17.2	9.0	8.3, 9.7
Non-Hispanic	7.6	16.0	8.4	8.0, 8.8
Unknown/missing	8.5	18.1	9.6	7.6, 11.7

eTable4 Difference in modeled proportions with low birthweight (<2500g) associated with, shift in means of 50, 125g, 167g, 250g and mean birthweight by social determinants in the ECHO consortium ordered according to mean birthweight (lowest to highest , N=28,496)

Category	N	Mean BW	Change in LBW (% points) Change in mean birthweight			
Category		Mean BW	50g	125g	167g	250g
Less than high school	1064	3190	1.79	4.81	6.66	10.69
High school	2283	3249	1.58	4.27	5.95	9.65
College, no degree	4985	3273	1.53	4.13	5.75	9.29
Missing/unknown	9430	3315	1.27	3.50	4.94	8.21
Bachelor’s degree	5557	3340	1.29	3.52	4.94	8.09
Master’s degree or higher	5177	3345	1.21	3.35	4.72	7.82
No HI/public HI	6781	3266	1.52	4.12	5.75	9.35
More than one	2916	3273	1.51	4.10	5.71	9.26
Missing/unknown	9056	3292	1.41	3.86	5.41	8.85
Employer/market/private	9743	3363	1.14	3.17	4.47	7.45
Black	4303	3131	1.99	5.32	7.36	11.76
Asian	993	3152	1.86	5.05	7.06	11.48
Other races	1295	3249	1.56	4.23	5.91	9.61
Multiple Races	3002	3271	1.51	4.09	5.71	9.27
Native Hawaiian/Pacific Islands	102	3285	1.41	3.86	5.41	8.89
Unknown/missing	1838	3309	1.28	3.55	5.00	8.30
American Indian/Alaskan Native	570	3356	1.27	3.47	4.86	7.94
White	16393	3373	1.12	3.11	4.40	7.31
Unknown/missing	872	3243	1.53	4.18	5.86	9.62
Hispanic	7174	3283	1.43	3.91	5.48	8.98
Non-Hispanic	20450	3319	1.33	3.64	5.10	8.37

eReferences

1. Peacock JL, Sauzet O, Ewings SM, Kerry SM. Dichotomising continuous data while retaining statistical power using a distributional approach. Stat Med 2012;31:3089-103.
2. Sauzet O, Peacock JL. Estimating dichotomised outcomes in two groups with unequal variances: a distributional approach. Stat Med 2014;33:4547-59.
3. Sauzet O, Ofuya M, Peacock JL. Dichotomisation using a distributional approach when the outcome is skewed. BMC Med Res Methodol 2015;15:40.
4. Sauzet O, Breckenkamp J, Borde T, et al. A distributional approach to obtain adjusted comparisons of proportions of a population at risk. Emerg Themes Epidemiol 2016;13:8.
5. Sauzet OK, Maren. Distributional estimates for the comparison of proportions of a dichotomized continuous outcome. The Stata Journal 2016;16:880-99.
6. DistdichoR a R Package for the distributional dichotomisation of continuous outcomes. 2018. (Accessed 12/5/23, at https://arxiv.org/abs/1809.03279)