Supplemental Online Content

Welch BM, Keil AP, Buckley JP, et al; the Pooled Phthalate Exposure and Preterm Birth Study Group. Associations between prenatal urinary biomarkers of phthalate exposure and preterm birth: a pooled study of 16 US cohorts. *JAMA Pediatr*. Published online July 11, 2022. doi:10.1001/jamapediatrics.2022.2252

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods 1. Description of study identification and inclusion in the pooled analysis

We performed a comprehensive search of PubMed for articles published through May 13th, 2019 to identify potentially eligible cohorts. The following terms were used in independent PubMed searches: "phthalates and gestation"; "phthalates and gestational"; "phthalates and preterm"; and "phthalates and pregnancy." Abstracts and methods sections of articles were reviewed to determine eligibility. Cohorts were considered potentially eligible for inclusion in the pooled analysis if the article: was published in English (original or translation); used an epidemiologic study design; was conducted in the United States of America (USA) or a USA territory (e.g., Puerto Rico); enrolled women during or prior to pregnancy; gathered information about gestational age at delivery; and measured ≥1 urinary phthalate metabolite in maternal urine collected during pregnancy.

In total, we identified 21 unique pregnancy cohorts that fit these criteria. Our final inclusion criteria were that a study had >50 participants and responded to our data transfer requests. We excluded 4 studies due to participant sample sizes of ≤50[1-4] and 1 study due to no response from the corresponding author.[5] This provided a total of 16 eligible studies that were included in this pooled analysis. The study design for selecting studies and eligible participants is described in eFigure **1**.

eMethods 2. Description of data harmonization

2.1. Variables used to determine preterm birth

- *Gestational age at enrollment and delivery*. Gestational age at enrollment and delivery was provided by all studies and converted to completed weeks (to first decimal) if not already provided as such. EPS participants were recruited before pregnancy so all gestational age at enrollment was set to "0." HEBC participants did not have a gestational age for first urine collection provided, but the value was set to 10 weeks based on the reported median value.[6] As detailed in Table 1, gestational age was defined by last menstrual period, early pregnancy ultrasound, date of conception in pregnancies utilizing assisted reproductive technologies (ARTs), or some combination thereof.
- *Preterm birth*. Preterm birth was defined as having a gestational age at delivery of <37 weeks, while term birth was ≥37 weeks gestation.

2.2. Variables used to assess phthalate exposures

- *Limit of detection (LOD) flags for phthalate biomarker concentrations.* Studies provided variables with specific LOD values for each biomarker measurement. Additionally, variables were provided or generated that flagged concentrations based on the LOD value, including the following categories: At or Above LOD; Below LOD-Instrument-Read Value; Below LOD-Imputed; Below LOD-Other (Reported as N/A, unknown, or 0); and Missing. Any concentrations below the LOD, but not explicitly stated as being an instrument-read value, were subsequently imputed as described in eMethods part C. Missing biomarker concentrations were not altered.
- *Urine specific gravity (SG) and creatinine.* Continuous values for SG and creatinine were provided for all studies.
- *Gestational age at urine collection.* This variable was reported in weeks and based on gestational age as described in eMethods B.1.

2.3. Variables used as primary confounders

- *Maternal race/ethnicity.* Categories of race/ethnicity were self-reported by participants of all studies, but a wide range of categories were reported. Thus, we generated a composite measure of self-identified categories that were combined to maximize sample size and consistency between pooled studies, including non-Hispanic White (Caucasian, White), non-Hispanic Black (African American, Black), Hispanic/Latina (Hispanic, Latino, Latin American indigenous heritage), Other (American Indian/Alaskan Native, Native Hawaiian, >1 racial identity).
- *Maternal education.* Maternal education was provided in different forms by studies. We summarized education to include the following categorical levels: less than high school (did not graduate); high school (graduated); some college (attended but did not graduate); college graduate (graduated undergraduate); graduate school. The "some college" category includes participants who reported attending some college or some technical school or 13-15 years of education. The "college graduate" category includes participants who reported receiving an undergraduate degree and/or attending ≥16 years of education. The "graduate school" category includes participants who reported receiving some graduate work or a graduate/advanced degree, as well as ≥17 years of education. Education information was not collected among HEBC participants,[6] but values were multiply imputed for the purposes of regression analyses.
- *Maternal age.* Maternal age was reported continuously for all studies except MSSM, which reported age as a categorical variable. The original categorical levels of maternal age among MSSM participants were: Less than 20; 20-<25; 25-<30; 30-<35; and ≥35; which we replaced with the continuous values 19, 22, 27, 32, and 37, respectively.

- *Maternal body mass index (BMI) –pre- and early pregnancy.* BMI values were reported as continuous values of kg/m^2. Prepregnancy BMI values were used whenever available, but early pregnancy values were used if prepregnancy values were unavailable (i.e., RDS). BMI measures were not available for SFF and Rutgers participants,[7, 8] but these values were multiply imputed for purposes of regression analyses.

2.4 Covariates used for descriptive statistics and/or as predictors in imputation models

- *Year of delivery.* The final variable of year measured on a continuous scale. For LIFECODES, TIDES, PROTECT, Healthy Start, RDS, MMIP, MSSM, EARTH, MARBLES, Rutgers, and SFF studies, year of delivery was available. For CHAMACOS, CCCEH, HOME, and EPS studies, year was abstracted based on year of urine collection which may differ in some pregnancies from year of delivery. For HEBC, a range of years was available from study notes from publications.[6, 9] For HEBC, the median of the year from the range was assigned to all the participants in that study. Additionally, there were 60 participants of SFF missing year of delivery, which was also imputed as 2002 based on the median from the range of years in the cohort (2000-2005).
- *Fetal sex.* Fetal sex was provided as male or female by all studies.
- *Parity*. Parity was recategorized as nulliparous or parous. A participant was categorized as parous if they reported having ≥1 prior pregnancy. Participants of MSSM were all nulliparous based on study design.[9]
- *Smoking.* A participant was categorized as "yes" for smoking in pregnancy if they reported ever smoking during pregnancy. Participants in HOME and CCCEH were categorized based on serum cotinine values, with "yes" defined by values ≥3 ng/mL.[10, 11]
- *Assisted Reproductive Technology (ART).* ART was categorized as "yes" if the participant reported using any of the following methods for the index pregnancy: IVF, ICSI, Donor Egg, or Other. ART was only used as a predictor in imputation models because it was only measured in a subset of studies (**eTable 3**).
- *Preeclampsia.* Dichotomous preeclampsia values (yes/no) were provided by all studies. Participants were reported as "no" if they reported "don't know", as was the case with CHAMACOS. Preeclampsia was only used as a predictor in imputation models.
- *Household income.* Income was only used as a predictor in imputation models because it was only measured in a subset of studies (**eTable 3**). Final income categories reflect household income and are coded into $10,000 range groupings (e.g. "Less than $10,000," "$10,000 - $19,999") until the household income exceeds $70,000. All incomes above $70,000 are grouped together (e.g., "$70,000+"). Original data from studies were in the form of income ranges. Additionally, the study dates for the different study ranged from 1983 to 2018. To account for the variability in reporting and collection times, we took the following steps. First, each participant was assigned their mean of the range of income. If income was reported as "$X or more", we retained the lowest income level within that range (e.g., if the range was "$150,000 or more," participants' income was coded as $150,000). Second, we account for inflation by calculating the inflation index for each study as of January 2020 using the Bureau of Labor Statistics Inflation Calculator. Inflation index is calculated using the original and current year. Original year was selected from the delivery year of each participant as described above. Third, we multiplied the income calculated in Step 1 with the inflation index calculated in Step 2. Fourth, using the adjusted income, household income was placed into $10,000 ranges. The lowest income level across the studies for the original "$X or more" ranges was $70,000, thus, for the current study, the highest category is "$70,000 or more."

eMethods 3. Multiple imputation

The goal for performing multiple imputation by chained equations (MICE) was to simultaneously impute: 1) phthalate metabolite concentrations below the limit of detection (LOD); and 2) missing covariate observations. We imputed phthalate metabolite concentrations below the LOD exclusively in the case where no instrument-read values were available (**eTable 5**). The proportions of samples with concentrations below the LOD that required imputation were relatively small across phthalate metabolites and ranged from 0.3% to 11% (**eTable 5**). Values below the LOD were imputed using a left-censored linear regression. The model assumed a log-normal distribution for each metabolite that was constrained to be between zero and the LOD, but allowed for the LOD value to vary within and across individuals (i.e., batch- and cohort-specific values). Missing covariate values were imputed by multivariate chained equations that used either predictive mean matching, logistic regression, or multinomial logistic regression for continuous, binary, and categorical covariates, respectively.[12] Primary covariates that were imputed included fetal sex (male; female) and the primary confounders of maternal age (years), race/ethnicity (non-Hispanic [NH] white; NH Black; Hispanic/Latina; Other), education (<high school; high school; some college; college graduate; graduate school), and prepregnancy body mass index (BMI).

Predictors used in MICE algorithms for concentrations below LOD and missing covariates included gestational age at delivery (weeks), gestational age at sample collection (weeks), study indicator (categorical, including sub-sites within study for TIDES and SFF), fetal sex (male; female), phthalate metabolite concentrations (continuous), and the previously listed set of confounders. Additionally, we included other covariates as predictors that were likely to be related to missing values, which is appropriate and improves accuracy of imputations when values may be missing not at random.[13] These predictors included preeclampsia (yes; no), parity (nulliparous; parous), smoking in pregnancy (yes; no), and use of assisted reproductive technology (yes; no).

We generated 10 imputed datasets using 20 chained iterations per dataset. Convergence of imputations was determined from trace plots of every imputed variable. Imputations were deemed to achieve adequate convergence based on minimal to no trends and strong mixing in concentrations across imputed data sets and iterations.[12] Imputation was carried out in R using MICE in the *mice* package (version 3.11.0).[13] Left-censored imputation of metabolite concentrations below LOD was done using the mice.impute.leftcenslognorm function from the *qgcomp* package (version 2.7.0).

eMethods 4. Methods to standardize phthalate metabolite concentrations by urine dilution

We implemented covariate-adjusted standardization to correct phthalate metabolite concentrations for urine dilution.[14, 15] This approach estimates a dilution-corrected value for each metabolite concentration. The covariate adjustment accounts for covariates that may influence hydration status (urinary specific gravity [SG] or creatinine), urinary phthalate metabolite concentrations, and/or an outcome of interest (i.e., preterm birth).[14] The method facilitates pooling data by allowing for comparisons of SG- and creatinine-standardized biomarker concentrations on the same scale.[15]

We fit cohort-specific models to generate fitted (covariate-adjusted) SG and creatinine values. In this study, we specified the following variables as relevant covariate predictors based on evidence in prior studies: maternal race/ethnicity, education, age, prepregnancy BMI, gestational age at urine sampling, and year of delivery.[15-18] Categorical covariates were included to account for studies with multiple study centers (TIDES and SFF). For example, SFF urine dilution model included a categorical variable indicating specific study-site locations in different states, including CA, MN, MO, and IA.[7] We used SG values if participants had both SG and creatinine values available (**eTable 1**): this included participants of the CHAMACOS and CCCEH studies.[10, 19]

For creatinine, we first fit a linear model for log-transformed creatinine concentrations as a function of the covariates, which is used to generate model-fitted values of creatinine for each participant.[14] These values were subsequently exponentiated to provide covariate-adjusted, or model-fitted, creatinine concentrations. We then created creatinine-standardized phthalate metabolite concentrations using the following formula: $E_{cor} = E_{obs} \times \frac{Cr_{fit}}{Cr_{obs}}$, where E_{cor} is the creatinine-standardized phthalate metabolite concentration, E_{obs} is the observed phthalate metabolite concentration, Cr_{fit} is the model-fitted creatinine concentration, and Cr_{obs} is the observed creatinine concentration. For SG, we used a modified version of this method previously established.[15] We first generated model-fitted values of SG for each participant by fitting a linear model for log-transformed SG as a function of the same covariate set. These values were subsequently exponentiated to provide covariate-adjusted, or model-fitted, SG values for every participant. We then created SG-standardized phthalate metabolite values using the following formula: $E_{cor} = E_{obs} \times \frac{SG_{fit}-1}{SG_{obs}-1}$, where E_{cor} is the SG-standardized phthalate metabolite concentration, E_{obs} is the observed phthalate metabolite concentration, SG_{fit} is the model-fitted SG value, SG_{obs} is the observed SG value. Since both the creatinine and SG approaches are based upon using the ratio of observed to fitted concentrations, the ratio measure is unitless. Thus, the resulting standardized phthalate metabolite concentrations from either method can be directly compared and are in the original units of the phthalate metabolite concentration (ng/mL).

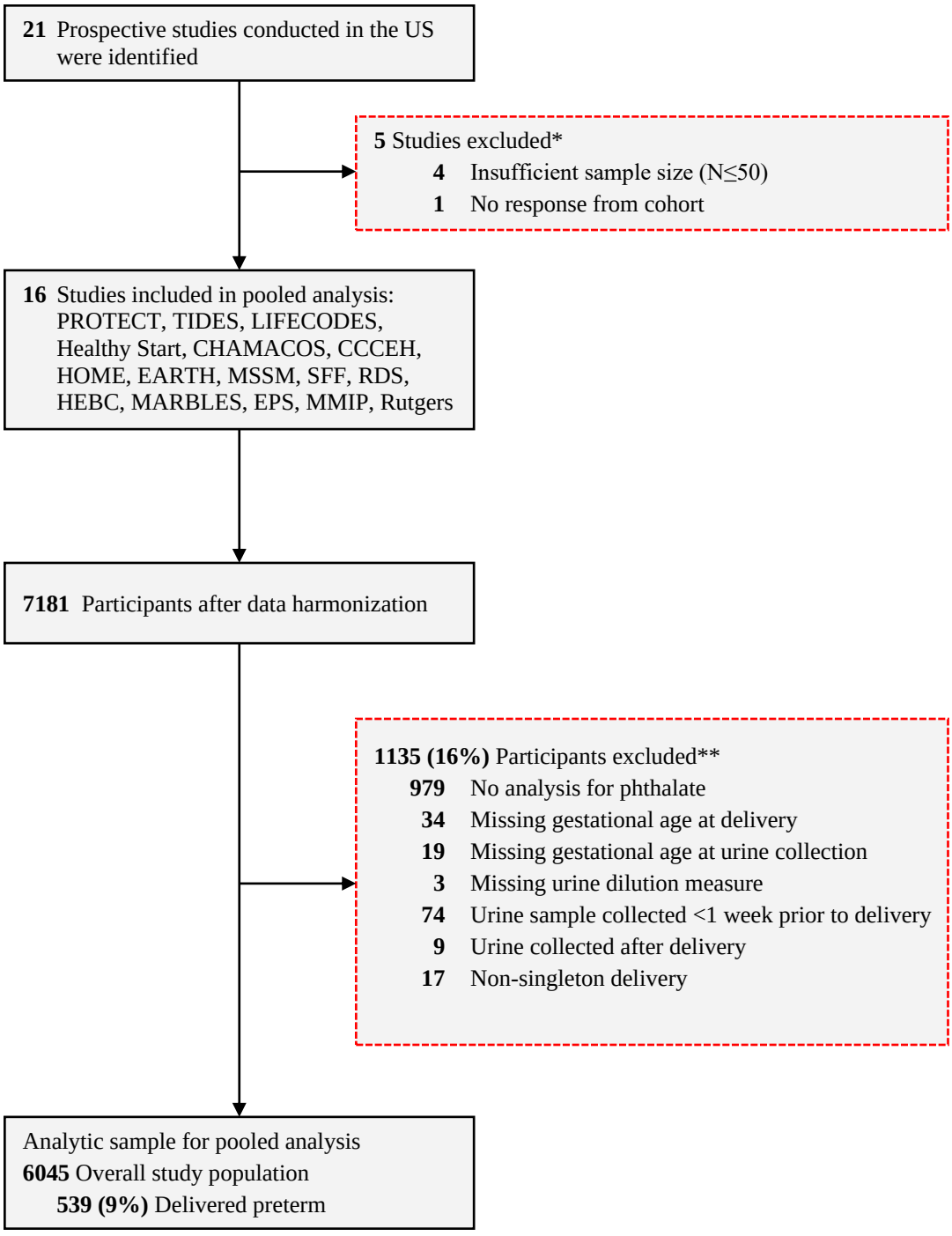
eMethods 5. Assumptions of g-computation necessary to infer causality

We used g-computation to determine potential changes in preterm birth following a range of hypothetical interventions that produced lower concentrations of a mixture of urinary phthalate metabolites within our pooled study population. The use of g-computation to evaluate hypothetical interventions is common for epidemiologic analyses in many subject areas, including environmental health,[20-23] as well as to improve interpretability of results or infer possible causal effects.[24] However, inferring possible causality requires a set of assumptions to be met. Within the context of exposure mixtures[20] and preterm birth, the more relevant assumptions include:

- *Correct model specification*. An assumption that our primary model correctly represents the true relationship between urinary phthalates and preterm birth. Given the results for individual metabolite models that included quadratic terms (eTable 12), our assumption of a linear scale was likely met. Although it is possible that metabolite by metabolite interactions were possible, including any such interactions would decrease the translatability of results, which was the primary goal for this g-computation analysis.

- *Exchangeability*. An assumption that there is no outstanding source of selection bias or confounding in our results. Within the context of our study, this assumption may be violated if selection bias was produced from phthalate exposure causing pregnancy to not result in a live birth.[25] However, given our study is principally interested in investigating associations among live births, it is unlikely to be a large source of bias. Another source of residual confounding could be diet, which can be a source of phthalate exposure[26] and risk factor for preterm birth.[27] Given phthalate exposure can come from many dietary pathways,[28] so the role of diet in is uncertain.
- *Positivity*. An assumption that there is a nonzero probability that phthalate metabolite concentrations can take on all possible values under the hypothetical interventions. This assumption is formally met within our analysis because phthalates can theoretically take on any nonnegative values, and we constrained phthalate concentrations from going below observed minimums. Our approach evaluated joint effects from simultaneously reducing all phthalate metabolites, which likely provides improved translatability to real-world exposure distributions.[20]
- *No measurement error of exposure*. An assumption that urinary phthalates were measured without systematic error. Variability in phthalate metabolite concentrations and use of single spot urine samples across certain studies may have been attributed to measurement error.
- *Treatment variation irrelevance*. An assumption that the effect of reducing phthalates via unspecified interventions will not product unanticipated impacts that adversely influence preterm birth. A relevant example may be that an intervention on one phthalate results in the substitution for another phthalate that also has an adverse influence on preterm birth. We recognize this assumption may not be fully achievable until the potential preterm birth effects of any such replacements are known.

eFigure 1. Flow diagram of study participant selection and exclusion in the Pooled Phthalate and Preterm Birth Study



Detailed description of study inclusion criteria provided in eMethods A and the exclusions by study are provided in eTable 2.

eTable 1. Additional study design elements of cohorts included in the Pooled Phthalate and Preterm Birth Study population

Study	Eligibility criteria	Recruitment sites	Type of urine sampling	Lab location and method	Urine dilution measure [a]	Urine samples per pregnancy average (med [min, max])
PROTECT	• Age 18-40 years • Residence within the Northern Karst aquifer region • Did not use oral contraceptives within the three months prior to pregnancy • No use of *in vitro* fertilization to become pregnant • No major preexisting medical conditions (e.g., diabetes)	Hospitals and health clinics in northern coast region of Puerto Rico	Spot	CDC[29]	SG	2 (1, 3)
TIDES	• Age ≥18 years • <13 weeks gestation • English speaking • No major pregnancy complications • Plans to deliver at participating hospital	Obstetrical medical centers at: 1) UCSF; 2) UMN; 3) URMC; and 4) SCH/UW	Spot	University of Washington[30] and CDC[29]	SG	2 (1, 3)
LIFECODES	• Non-anomalous fetus • Live singleton birth • Plans to delivery at BWH	Tertiary care clinics of Brigham Women's Hospital in Boston, Massachusetts	Spot	NSF International[29]	SG	4 (1, 4)
Healthy Start	• Age ≥16 years • <24 weeks gestation • No prior stillbirth, diabetes, asthma, cancer, or serious psychiatric illness	Obstetric clinics at the University of Colorado Hospital in Aurora, Colorado	Spot	CDC[29]	Creatinine	1 (1, 1)
CHAMACOS	• English or Spanish speaking • ≤20 weeks pregnant • ≥18 years old • Low income (Medi-Cal California Medicaid eligible) • Intention to deliver at county hospital	Six prenatal clinics serving farmworkers in Salinas Valley, California	Spot	CDC[29]	SG & creatinine	2 (1, 2)
CCCEH	• Age 18-35 years • First prenatal visit <20 weeks gestation • African American or Dominican identity • Living in northern Manhattan or South Bronx for ≥1 year prepregnancy • No tobacco or drug use in pregnancy No chronic medical conditions (HIV, diabetes, hypertension)	Prenatal clinics at Harlem and New York (NY) Presbyterian hospitals in NY City, NY	Spot	CDC[29]	SG & creatinine	1 (1, 1)

Study	Eligibility criteria	Recruitment sites	Type of urine sampling	Lab location and method	Urine dilution measure [a]	Urine samples per pregnancy average (med [min, max])
HOME	• Age ≥18 years • 16±3 weeks gestation • Living in surrounding counties and intention to deliver at participating clinics • Living in home (no mobile/trailer home) built ≤1978 (related to original focus on lead exposure) • No chronic medical conditions (HIV, diabetes, bipolar disorder, schizophrenia, chemotherapy- or radiation-treated cancer) • No genetic abnormalities or birth defects	Prenatal practices of three hospitals in region surrounding Cincinnati, Ohio	Spot	CDC[29]	Creatinine	2 (1, 2)
EARTH	• Age 18-46 years (women) • One prepregnancy urine sample taken prior to conception of index pregnancy (only pregnancy measures evaluated here)	Massachusetts General Hospital Fertility Center in Boston, Massachusetts	Spot	CDC[29]	SG	3 (1, 3)
MSSM	• Primiparous (first pregnancy/nulliparous) • No chronic conditions (diabetes, hypertension, thyroid disease) • No serious pregnancy complications (delivery <32 weeks, or fetal genetic abnormalities or malformations) or change in residence outside NY City	Prenatal clinic and private practices at Mount Sinai Medical Center in NY City, NY	Spot	CDC[31]	Creatinine	1 (1, 1)
SFF	• Age ≥18 years • Natural conception • No severe pregnancy complications • Live within 50 miles of clinic Participated in postpartum follow-up study	Prenatal clinics of university hospitals in: 1) Los Angeles, California; 2) Minneapolis, Minnesota; 3) Columbia, Missouri; and 4) Iowa City, Iowa	Spot	CDC[32]	Creatinine	1 (1, 1)
RDS	• Age ≥18 years • First trimester ultrasound confirmed pregnancy • No fetal genetic anomalies or aneuploidy • No use of progesterone or other steroids • No chronic medical conditions (diabetes, thyroid or other endocrine disorder)	Medical University of South Carolina in metropolitan area of Charleston, South Carolina	Spot	National Institute of Standards and Technology, Charleston, South Carolina[32]	SG	1 (1, 2)
HEBC	• Women participated in prior enrollment studies and contributed first-trimester urine sample between 2007-2009 • No chronic medical conditions (diabetes, chronic hypertension)	Clinics and private practices affiliated with the Brigham and Women's Hospital in Boston, Massachusetts	Spot	CDC[29]	SG	1 (1, 1)

Study	Eligibility criteria	Recruitment sites	Type of urine sampling	Lab location and method	Urine dilution measure [a]	Urine samples per pregnancy average (med [min, max])
MARBLES	• High risk of delivering child who will develop autism spectrum disorder (ASD), primarily because previously delivered child who developed ASD • Age ≥18 years • English fluency • Lives within 2.5 hours of Davis/Sacramento region	Recruitment occurred primarily through California Department of Developmental Services, along with other sources (other studies, provider referrals), in Northern California	First morning void or 24 hour	CDC[29]	SG	3 (1, 10)*
EPS	• No diagnosed fertility problems • No chronic medical conditions	Recruitment via community advertisements in North Carolina	Pooled urine sample (3 samples collected over 3-week period)	CDC[29]	Creatinine	1 (1, 1)*
MMIP	• Age ≥18 years • Naturally conceived	Recruitment occurred during first prenatal visit at University of Michigan OG/GYN facility in Ann Arbor, Michigan	Spot	NSF International[33]	SG	1 (1, 1)
Rutgers	Age ≥18 years	Recruited from the High-Risk Obstetric Clinic at Robert Wood Johnson University Hospital, part of Rutgers University, in New Brunswick, New Jersey	Spot	Rutgers University[8]	SG	1 (1, 1)

Abbreviations: SG, specific gravity; med, median

[a] If both SG and creatinine were available, only SG was used.

* EPS and MARBLES combined (pooled) repeated urine samples together prior to measuring phthalate metabolites.

eTable 2. Description of participant exclusions and final sample size in the Pooled Phthalate and Preterm Birth Study population

	Original sample		Reason for exclusion (n)							Pooled analysis
	N	Excluded[a] (n [%])	No analysis for phthalates[b]	Missing gestational age at delivery	Missing gestational age at urine collection[c]	Missing urine dilution measure	Urine collected <1 week prior to delivery	Urine collected after delivery	Non-singleton delivery	Analytic sample (N)
Overall	7181	1136 (16)	979	34	19	3	74	9	17	6045
PROTECT	1128	27 (2)	24	0	0	0	3	0	0	1101
TIDES	969	190 (20)	187	0	0	0	1	0	2	779
LIFECODES	482	2 (0)	0	2	0	0	0	0	0	480
Healthy Start	446	2 (0)	0	2	0	0	0	0	0	444
CHAMACOS	596	167 (28)	167	0	0	0	0	0	0	429
CCCEH	456	67 (15)	0	29	4	0	29	5	0	389
HOME	389	0 (0)	0	0	0	0	0	0	0	389
EARTH	386	1 (0)	0	0	0	0	1	0	0	385
MSSM	404	42 (10)	22	0	3	1	12	4	0	362
SFF	955	602 (63)	575	1	0	0	20	0	6	353
RDS	319	1 (0)	0	0	0	0	1	0	0	318
HEBC	195	6 (3)	0	0	0	0	0	0	6	189
MARBLES	186	7 (4)	0	0	0	0	4	0	3	179
EPS	130	4 (3)	3	0	0	1	0	0	0	126
MMIP	68	0 (0)	0	0	0	0	0	0	0	68
Rutgers	72	18 (25)	1	0	12	1	4	0	0	54

[a] Study-specific percent value provided

[b] If all phthalate metabolite concentrations were missing for a participant, it was assumed that no urine samples were collected during pregnancy.

[c] Participants were excluded if gestational age at urine collection was missing because it was possible collection could have occurred <1 prior to delivery.

eTable 3. Participant characteristics (n [%] or mean [SD]) by study (a-p)

	a. PROTECT	b. TIDES	c. LIFECODES [a]	d. Healthy Start	e. CHAMACOS	f. CCCEH
Sample size (n)	1101	779	480	444	429	389
Delivery (n)						
Term	1001 (90.9)	710 (91.1)	350 (72.9)	430 (96.8)	402 (93.7)	375 (96.4)
Preterm	100 (9.1)	69 (8.9)	130 (27.1)	14 (3.2)	27 (6.3)	14 (3.6)
Gestational age at delivery (weeks)	38.9 (2.0)	39.3 (1.8)	38.0 (2.8)	39.5 (1.3)	39.0 (1.8)	39.3 (1.3)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal age (years)	27.1 (5.5)	31.0 (5.5)	32.1 (5.4)	28.2 (6.1)	26.8 (5.3)	25.3 (4.8)
Missing	1 (0.1)	2 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal race/ethnicity (n)						
Non-Hispanic White	0 (0.0)	511 (65.6)	283 (59.0)	255 (57.4)	7 (1.6)	0 (0.0)
Non-Hispanic Black	0 (0.0)	95 (12.2)	76 (15.8)	49 (11.0)	0 (0.0)	132 (33.9)
Hispanic/Latina	1101 (100.0)	68 (8.7)	71 (14.8)	109 (24.5)	414 (96.5)	257 (66.1)
Other	0 (0.0)	96 (12.3)	50 (10.4)	31 (7.0)	8 (1.9)	0 (0.0)
Missing	0 (0.0)	9 (1.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal education (n)						
Less than high school	228 (20.7)	61 (7.8)	17 (3.5)	60 (13.5)	337 (78.6)	147 (37.8)
High school	108 (9.8)	48 (6.2)	49 (10.2)	71 (16.0)	49 (11.4)	139 (35.7)
Some college	602 (54.7)	95 (12.2)	73 (15.2)	98 (22.1)	18 (4.2)	69 (17.7)
College graduate	119 (10.8)	240 (30.8)	143 (29.8)	100 (22.5)	25 (5.8)	30 (7.7)
Graduate school	26 (2.4)	326 (41.8)	187 (39.0)	115 (25.9)	0 (0.0)	4 (1.0)
Missing	18 (1.6)	9 (1.2)	11 (2.3)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal prepregnancy BMI (kg/m^2)	25.3 (5.5)	25.7 (6.4)	25.8 (6.0)	25.7 (6.4)	27.2 (5.3)	25.6 (5.9)
Missing	61 (5.5)	7 (0.9)	0 (0.0)	0 (0.0)	9 (2.1)	6 (1.5)
Delivery year (n)						
1983-2000	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	377 (87.9)	102 (26.2)
2001-2010	0 (0.0)	0 (0.0)	480 (100.0)	0 (0.0)	52 (12.1)	287 (73.8)
2011-2018	1101 (100.0)	779 (100.0)	0 (0.0)	444 (100.0)	0 (0.0)	0 (0.0)
Maternal smoking during pregnancy (n)						
No	1069 (97.1)	718 (92.2)	452 (94.2)	412 (92.8)	406 (94.6)	332 (85.3)
Yes	17 (1.5)	57 (7.3)	28 (5.8)	32 (7.2)	23 (5.4)	8 (2.1)
Missing	15 (1.4)	4 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	49 (12.6)
Fetal sex (n)						
Female	516 (46.9)	397 (51.0)	213 (44.4)	204 (45.9)	211 (49.2)	203 (52.2)
Male	579 (52.6)	382 (49.0)	267 (55.6)	240 (54.1)	218 (50.8)	186 (47.8)
Missing	6 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Parity (n)						
Nulliparous	538 (48.9)	393 (50.4)	214 (44.6)	220 (49.5)	142 (33.1)	179 (46.0)
Parous	550 (50.0)	332 (42.6)	266 (55.4)	224 (50.5)	287 (66.9)	209 (53.7)
Missing	13 (1.2)	54 (6.9)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.3)

	g. HOME	h. EARTH	i. MSSM [b]	j. SFF	k. RDS	l. HEBC
Sample size (n)	389	385	362	353	318	189

Delivery (n)						
Term	352 (90.5)	358 (93.0)	334 (92.3)	336 (95.2)	290 (91.2)	177 (93.7)
Preterm	37 (9.5)	27 (7.0)	28 (7.7)	17 (4.8)	28 (8.8)	12 (6.3)
Gestational age at delivery (weeks)	39.0 (1.8)	39.4 (1.7)	39.3 (1.6)	39.3 (1.6)	38.8 (1.8)	38.9 (1.3)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal age (years)	29.3 (5.8)	34.7 (3.9)	23.9 (5.6)	30.2 (5.1)	27.7 (5.6)	32.9 (5.1)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	13 (3.7)	0 (0.0)	0 (0.0)
Maternal race/ethnicity (n)						
Non-Hispanic White	237 (60.9)	327 (84.9)	76 (21.0)	296 (83.9)	158 (49.7)	133 (70.4)
Non-Hispanic Black	120 (30.8)	11 (2.9)	107 (29.6)	6 (1.7)	151 (47.5)	23 (12.2)
Hispanic/Latina	9 (2.3)	0 (0.0)	178 (49.2)	31 (8.8)	3 (0.9)	26 (13.8)
Other	18 (4.6)	47 (12.2)	1 (0.3)	18 (5.1)	6 (1.9)	7 (3.7)
Missing	5 (1.3)	0 (0.0)	0 (0.0)	2 (0.6)	0 (0.0)	0 (0.0)
Maternal education (n)						
Less than high school	41 (10.5)	0 (0.0)	104 (28.7)	7 (2.0)	30 (9.4)	0 (0.0)
High school	54 (13.9)	0 (0.0)	76 (21.0)	19 (5.4)	57 (17.9)	0 (0.0)
Some college	93 (23.9)	0 (0.0)	94 (26.0)	72 (20.4)	79 (24.8)	0 (0.0)
College graduate	115 (29.6)	127 (33.0)	0 (0.0)	134 (38.0)	85 (26.7)	0 (0.0)
Graduate school	81 (20.8)	207 (53.8)	0 (0.0)	120 (34.0)	49 (15.4)	0 (0.0)
Missing	5 (1.3)	51 (13.2)	88 (24.3)	1 (0.3)	18 (5.7)	189 (100.0)
Maternal prepregnancy BMI (kg/m^2)	26.6 (6.5)	24.2 (4.3)	23.5 (4.5)	NA	29.2 (7.1)	25.5 (6.0)
Missing	0 (0.0)	0 (0.0)	1 (0.3)	353 (100.0)	1 (0.3)	1 (0.5)
Delivery year (n)						
1983-2000	0 (0.0)	0 (0.0)	310 (85.6)	4 (1.1)	0 (0.0)	0 (0.0)
2001-2010	389 (100.0)	144 (37.4)	52 (14.4)	349 (98.9)	0 (0.0)	189 (100.0)
2011-2018	0 (0.0)	241 (62.6)	0 (0.0)	0 (0.0)	318 (100.0)	0 (0.0)
Maternal smoking during pregnancy (n)						
No	335 (86.1)	289 (75.1)	300 (82.9)	339 (96.0)	276 (86.8)	183 (96.8)
Yes	53 (13.6)	96 (24.9)	62 (17.1)	13 (3.7)	39 (12.3)	6 (3.2)
Missing	1 (0.3)	0 (0.0)	0 (0.0)	1 (0.3)	3 (0.9)	0 (0.0)
Fetal sex (n)						
Female	208 (53.5)	185 (48.1)	163 (45.0)	146 (41.4)	133 (41.8)	99 (52.4)
Male	181 (46.5)	200 (51.9)	199 (55.0)	150 (42.5)	185 (58.2)	87 (46.0)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	57 (16.1)	0 (0.0)	3 (1.6)
Parity (n)						
Nulliparous	171 (44.0)	320 (83.1)	362 (100.0)	187 (53.0)	128 (40.3)	71 (37.6)
Parous	216 (55.5)	65 (16.9)	0 (0.0)	165 (46.7)	190 (59.7)	117 (61.9)
Missing	2 (0.5)	0 (0.0)	0 (0.0)	1 (0.3)	0 (0.0)	1 (0.5)

	m. MARBLES	n. EPS	o. MMIP	p. Rutgers
Sample size (n)	179	126	68	54
Delivery (n)				
Term	167 (93.3)	121 (96.0)	66 (97.1)	37 (68.5)
Preterm	12 (6.7)	5 (4.0)	2 (2.9)	17 (31.5)
Gestational age at delivery (weeks)	38.9 (1.6)	40.0 (1.8)	39.6 (1.1)	37.6 (2.5)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal age (years)	34.0 (5.0)	29.0 (3.6)	31.7 (4.6)	33.2 (6.6)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal race/ethnicity (n)				
Non-Hispanic White	99 (55.3)	120 (95.2)	56 (82.4)	18 (33.3)
Non-Hispanic Black	10 (5.6)	3 (2.4)	4 (5.9)	15 (27.8)
Hispanic/Latina	38 (21.2)	0 (0.0)	2 (2.9)	16 (29.6)
Other	32 (17.9)	3 (2.4)	6 (8.8)	5 (9.3)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal education (n)				
Less than high school	5 (2.8)	0 (0.0)	0 (0.0)	8 (14.8)
High school	8 (4.5)	9 (7.1)	5 (7.4)	14 (25.9)
Some college	72 (40.2)	26 (20.6)	6 (8.8)	13 (24.1)
College graduate	69 (38.5)	46 (36.5)	18 (26.5)	12 (22.2)
Graduate school	25 (14.0)	45 (35.7)	31 (45.6)	7 (13.0)
Missing	0 (0.0)	0 (0.0)	8 (11.8)	0 (0.0)
Maternal prepregnancy BMI (kg/m^2)	26.8 (6.9)	21.1 (2.8)	25.4 (5.5)	NA
Missing	0 (0.0)	0 (0.0)	3 (4.4)	54 (100.0)
Delivery year (n)				
1983-2000	0 (0.0)	126 (100.0)	0 (0.0)	0 (0.0)
2001-2010	115 (64.2)	0 (0.0)	2 (2.9)	54 (100.0)
2011-2018	64 (35.8)	0 (0.0)	66 (97.1)	0 (0.0)
Maternal smoking during pregnancy (n)				
No	161 (89.9)	120 (95.2)	65 (95.6)	42 (77.8)
Yes	8 (4.5)	6 (4.8)	3 (4.4)	12 (22.2)
Missing	10 (5.6)	0 (0.0)	0 (0.0)	0 (0.0)
Fetal sex (n)				
Female	75 (41.9)	59 (46.8)	32 (47.1)	26 (48.1)
Male	104 (58.1)	67 (53.2)	36 (52.9)	28 (51.9)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Parity (n)				
Nulliparous	2 (1.1)	60 (47.6)	31 (45.6)	9 (16.7)
Parous	171 (95.5)	66 (52.4)	37 (54.4)	45 (83.3)
Missing	6 (3.4)	0 (0.0)	0 (0.0)	0 (0.0)

SD, standard deviation; BMI, body mass index; ART, assisted reproductive technology; NA, not assessed

[a] LIFECODES was a case-control study of preterm birth; [b] Year of delivery was assigned as the median year, 2000 (see eMethods)

eTable 4. Urinary metabolites of phthalate and phthalate alternative compounds measured in the Pooled Phthalate and Preterm Birth study

						Pooled sample		
Parent chemical [a]			Metabolite			Cohort	Participants	Analysis [b]
Name	Abbrev.	MW	Name	Abbrev.	MW	(16 total)	(N=6045)	
Dimethyl-phthalate	DMP	194	Monomethyl phthalate	MMP	180	5	23%	Excluded
Diethyl phthalate	DEP	222	Monoethyl phthalate	MEP	194	16	100%	**Included**
Di-n-butyl phthalate	DBP	278	Mono-n-butyl phthalate	MBP	222	16	100%	**Included**
			Mono(3-hydroxybutyl) phthalate	MHBP	238	4	24%	Excluded
Di-isobutyl phthalate	DiBP	278	Mono-isobutyl phthalate	MiBP	222	15	99%	**Included**
			Mono-hydroxyisobutyl phthalate	MHiBP	238	4	24%	Excluded
Benzylbutyl phthalate	BzBP	312	Monobenzyl phthalate	MBzP	256	15	99%	**Included**
Dicyclohexyl phthalate	DCHP	330	Mono-cyclohexyl phthalate	MCHP	248	1	1%	Excluded
Di(2-ethylhexyl) phthalate	DEHP	391	Mono(2-ethylhexyl) phthalate	MEHP	278	16	97%	**Included**
			Mono(2-ethyl-5-hydroxyhexyl) phthalate	MEHHP	294	16	100%	**Included**
			Mono(2-ethyl-5-carboxypentyl) phthalate	MECPP	308	14	91%	**Included**
			Mono(2-ethyl-5-oxyhexyl) phthalate	MEOHP	422	16	100%	**Included**
Di(2-ethylhexyl) terephthalate	DEHTP	391	Mono(2-ethyl-5-hydroxyhexyl) terephthalate	MEHHTP	294	2	7%	Excluded
			Mono(2-ethyl-5-carboxypentyl) terephthalate	MECPTP	308	2	7%	Excluded
Di-n-octyl phthalate (and other high MW phthalates)	DNOP	391	Mono(3-carboxypropyl) phthalate	MCPP	252	14	93%	**Included**
			Mono-n-octyl phthalate	MOP	278	1	1%	Excluded
Di-isononyl phthalate	DNP	419	Monoisononyl phthalate	MNP	292	5	24%	Excluded
			Monooxoisononyl phthalate	MONP	292	2	7%	Excluded
			Monocarboxy-isooctyl phthalate	MCOP	322	10	57%	**Included**
1,2-Cyclohexane dicarboxylic acid, diisononyl ester	DINCH	425	Monocarboxy-isooctyl ester, 1,2-cyclohexane-dicarboxylic acid	MCOCH	172	4	19%	Excluded
			Monohydroxy-isononyl ester, 1,2-cyclohexane dicarboxylic acid	MHiNCH	314	5	38%	Excluded
Di-isodecyl phthalate	DDP	447	Monocarboxy-isononyl phthalate	MCNP	322	10	58%	**Included**
			Monoisodecyl phthalate	MDP	306	1	1%	Excluded

[a] Parent compounds are ordered by molecular weight (MW; g/mol).

[b] Analysis decision identifies whether metabolite was included or excluded from primary analyses. A given metabolite was included if it was measured in ≥10 cohorts and ≥50% of all participant samples.

eTable 5. Limits of detection (LOD) for phthalate metabolites and distribution of samples with concentrations above and below LOD

Biomarker	LOD range [a] (ng/ml)	Number of observations	% >LOD	% <LOD	% <LOD with instrument-read values [b]	% <LOD without instrument-read values [c]
MEP	0.40 - 1.20	11391	99.5%	0.5%	0.2%	0.3%
MBP	0.10 - 2.00	11391	98.3%	1.6%	0.3%	1.3%
MiBP	0.10 - 1.04	11337	97.3%	2.7%	0.7%	2.0%
MBzP	0.10 - 1.00	11337	96.2%	3.9%	1.3%	2.6%
MEHP	0.05 - 1.20	11391	82.5%	17.5%	6.5%	11.0%
MEHHP	0.10 - 1.00	11391	99.2%	0.8%	0.2%	0.6%
MECPP	0.20 - 1.00	10672	99.9%	0.1%	0.0%	0.1%
MEOHP	0.10 - 1.07	11391	99.1%	0.9%	0.1%	0.8%
MCPP	0.16 - 1.00	10874	90.3%	9.7%	4.2%	5.5%
MCOP	0.20 - 0.70	7094	99.0%	1.1%	0.4%	0.7%
MCNP	0.20 - 0.60	7130	97.0%	3.0%	0.8%	2.2%

[a] LOD is presented as a range because of variation across studies.

[b] Instrument-read values were used when available for concentrations <LOD .

[c] Concentrations <LOD were multiply imputed when instrument-read values were not available.

eTable 6. Sample size for each urinary phthalate metabolite across studies

Biomarkers[a]	MEP	MBP	MiBP	MBzP	MEHP	MEHHP	MECPP	MEOHP	MCPP	MCOP	MCNP
Overall											
Cohorts	16	16	15	15	16	16	14	16	14	10	10
Sample size	6045	6045	5991	5991	6045	6045	5471	6045	5673	3758	3794
Study											
PROTECT	1101	1101	1101	1101	1101	1101	1101	1101	1101	1101	1101
TIDES	779	779	779	779	779	779	779	779	779	754	754
LIFECODES	480	480	480	480	480	480	480	480	480	NM	NM
Healthy Start	444	444	444	444	444	444	444	444	444	444	444
CHAMACOS	429	429	429	429	429	429	429	429	429	429	429
CCCEH	389	389	389	389	389	389	389	389	389	146	146
HOME	389	389	389	389	389	389	389	389	389	NM	NM
EARTH	385	385	385	385	385	385	385	385	385	372	372
MSSM	362	362	362	362	362	362	362	362	362	NM	NM
SFF	353	353	353	353	353	353	151	353	353	18	54
RDS	318	318	318	318	318	318	NM	318	NM	NM	NM
HEBC	189	189	189	189	189	189	189	189	189	189	189
MARBLES	179	179	179	179	179	179	179	179	179	179	179
EPS	126	126	126	126	126	126	126	126	126	126	126
MMIP	68	68	68	68	68	68	68	68	68	NM	NM
Rutgers	54	54	NM	NM	54	54	NM	54	NM	NM	NM

NM, not measured

[a] The only biomarkers excluded from mixtures analyses were MCOP and MCNP.

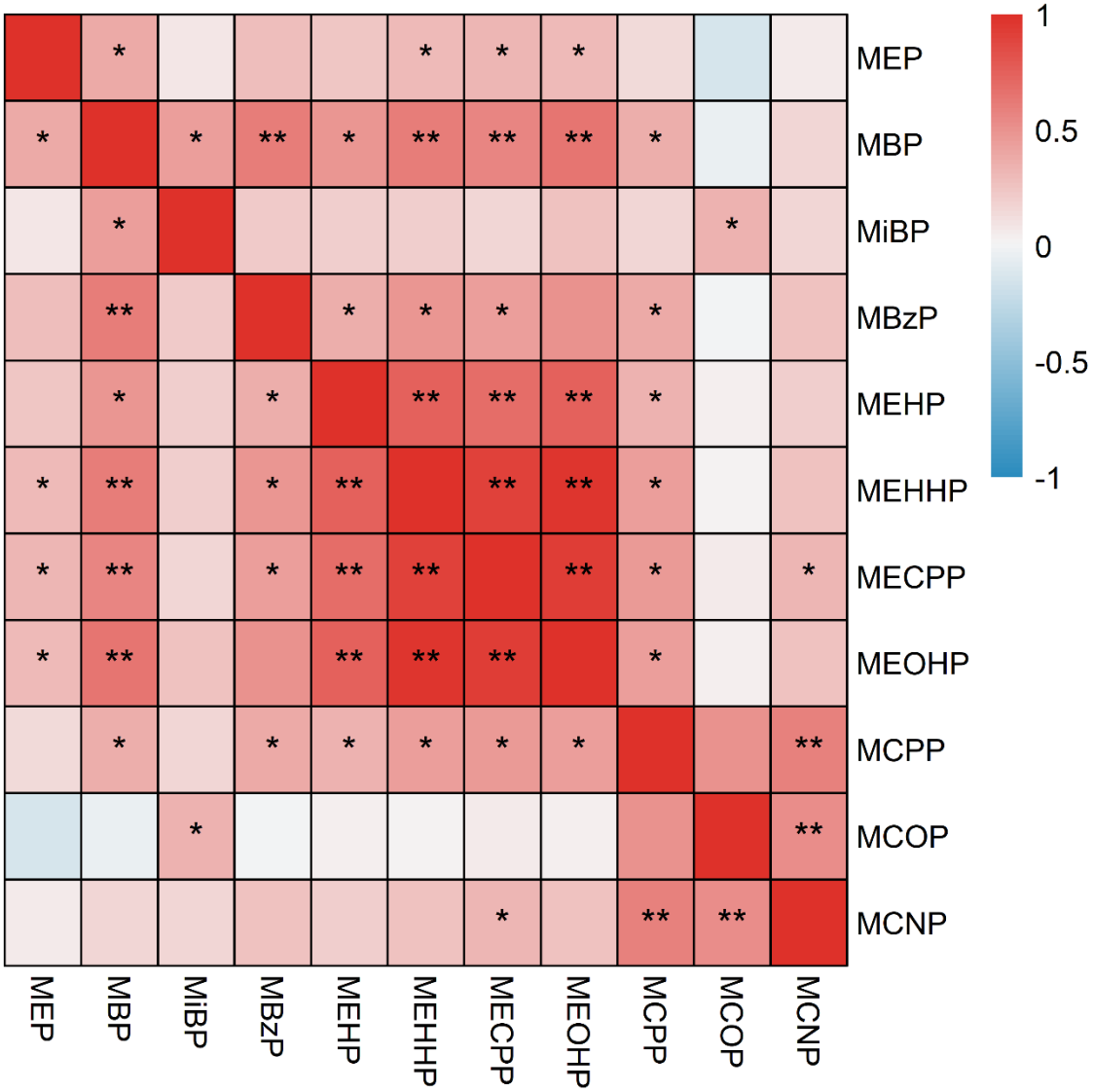
eTable 7. Distribution of pregnancy-averaged urinary phthalate metabolite concentrations (ng/mL)

Metabolite[a]	Cohorts	Sample size	GM	25th percentile	Median	75th percentile	IQR
MEP	16	6045	73.04	25.0	68.9	193.2	168.2
MBP	16	6045	16.06	8.7	15.5	30.1	21.4
MiBP	15	5991	6.16	3.3	6.3	11.9	8.6
MBzP	15	5991	5.93	2.5	5.6	13.4	11.0
MEHP	16	6045	3.12	1.5	2.9	6.4	5.0
MEHHP	16	6045	11.96	5.8	10.9	23.0	17.3
MECPP	14	5471	20.63	10.2	18.8	37.0	26.8
MEOHP	16	6045	9.29	4.7	8.6	17.1	12.4
MCPP	14	5673	2.05	1.1	1.9	3.6	2.5
MCOP	10	3758	10.06	4.1	9.2	22.7	18.5
MCNP	10	3794	2.36	1.4	2.2	3.6	2.2

GM, geometric mean; IQR, interquartile range

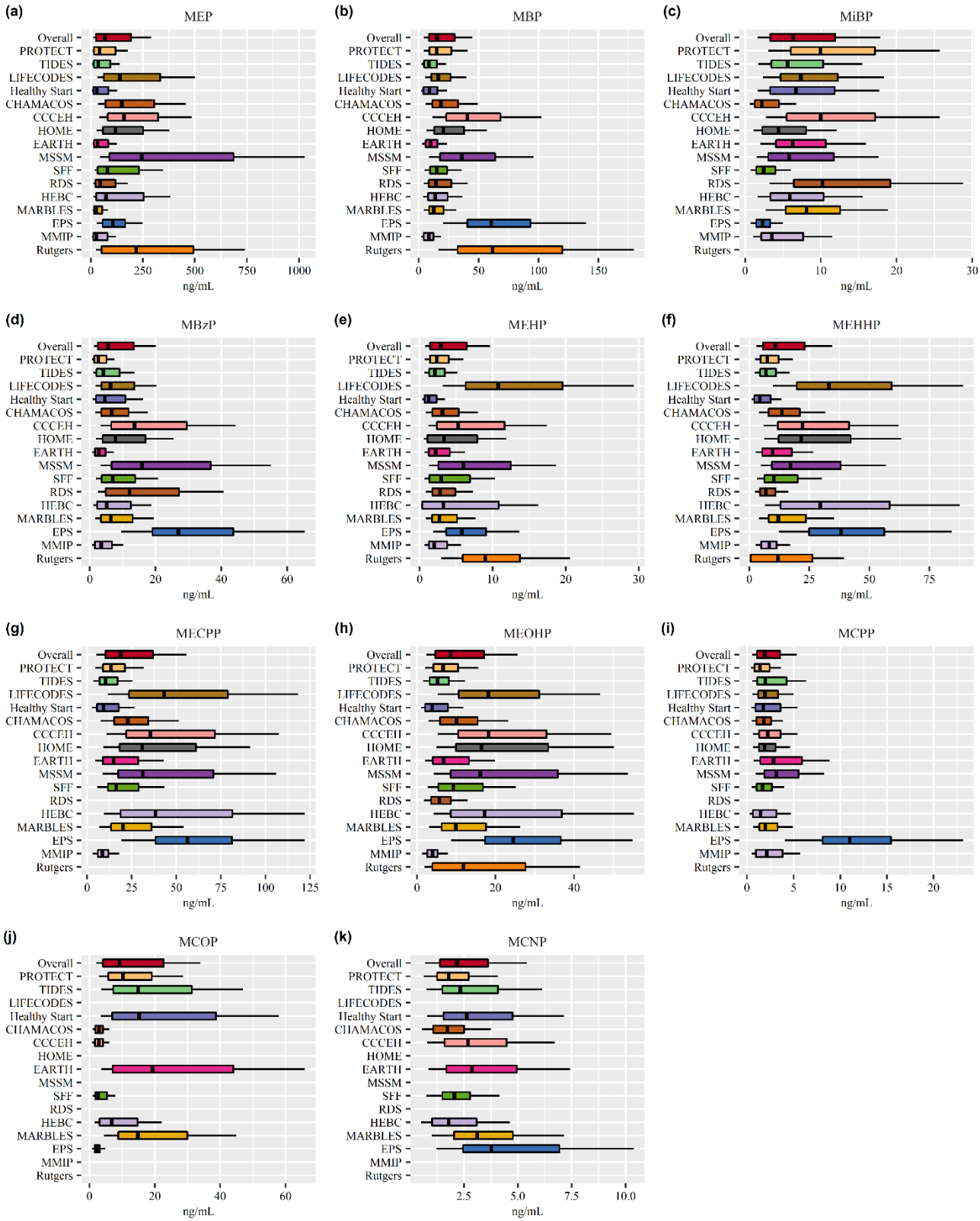
[a] Biomarker concentrations were corrected for urine dilution before pregnancy-averages were calculated; thus, all values are corrected for urine dilution.

eFigure 2. Spearman correlations between pregnancy-averaged concentrations of urinary phthalate metabolites



Asterisks indicate absolute correlation values between 0.3 and 0.5 (*), or greater than 0.50 (**).

eFigure 3. Distributions of pregnancy-averaged phthalate metabolite concentrations (a-k) in the Pooled Phthalate and Preterm Birth Study (overall) and by study



Concentrations were standardized by urine dilution. Each box shows the 25th, 50th, and 75th percentiles. The upper whisker represents 1.5 times the 75th percentile while the lower whisker represents 0.5 times the 25th percentile, stopping at the limit of detection. Values above or below whiskers not shown. Studies are ordered by the relative size of the study population.

eTable 8. Heterogeneity by study in main effects using fixed effect, random effect, and interaction models

		Fixed effect [a,b]		Random effect [a,c]		Heterogeneity in main effect (Wald test) [d]
Metabolite	n	OR (95% CI)	Variance	OR (95% CI)	Variance	Study*Metabolite
MEP	6045	1.07 (0.93,1.24)	0.0051	1.08 (0.94,1.24)	0.0050	0.35
MBP	6045	1.12 (0.98,1.27)	0.0045	1.13 (0.99,1.28)	0.0044	0.17
MiBP	5991	1.16 (1.00,1.34)	0.0058	1.17 (1.01,1.35)	0.0055	0.85
MBzP	5991	0.98 (0.83,1.14)	0.0065	0.97 (0.83,1.13)	0.0062	0.62
MEHP	6045	1.04 (0.91,1.19)	0.0048	1.06 (0.93,1.21)	0.0046	0.06
MEHHP	6045	1.03 (0.90,1.19)	0.0049	1.04 (0.91,1.18)	0.0047	0.28
MECPP	5471	1.16 (1.00,1.34)	0.0056	1.17 (1.01,1.35)	0.0053	0.54
MEOHP	6045	1.00 (0.88,1.15)	0.0046	1.01 (0.89,1.15)	0.0044	0.32
MCPP	5673	1.14 (1.01,1.29)	0.0039	1.13 (1.00,1.28)	0.0037	0.39
MCOP	3758	1.04 (0.84,1.29)	0.0119	1.08 (0.88,1.32)	0.0107	0.35
MCNP	3794	1.06 (0.92,1.24)	0.0059	1.05 (0.91,1.22)	0.0057	0.63

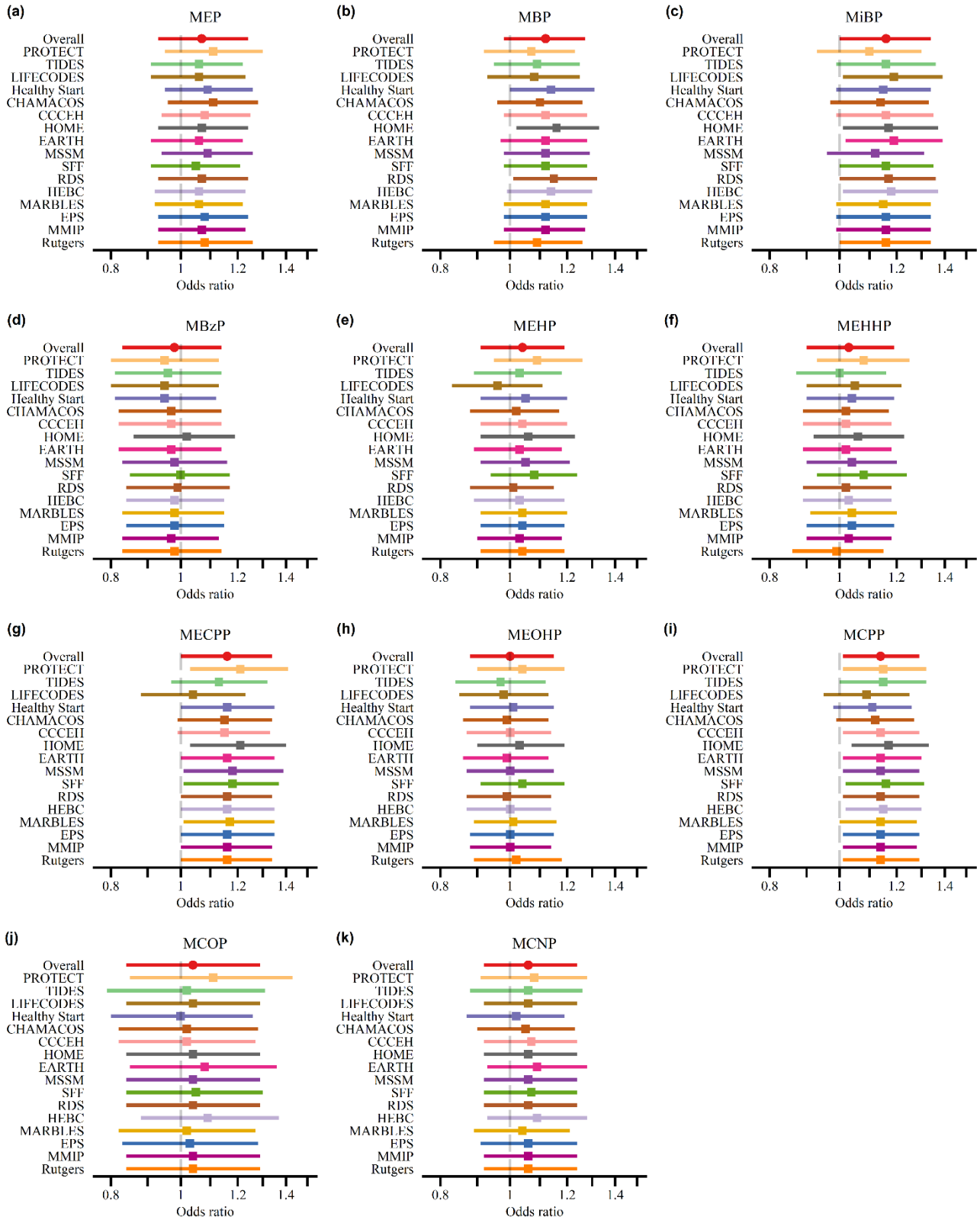
[a] OR and 95% confidence interval (CI) represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual biomarker. Associations were estimated by multiple logistic regression models. All models adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI. Variance estimates represent the standard error squared from non-transformed model estimates.

[b] Fixed effect adjusted models include study cohort as fixed effect covariate.

[c] Random effect models included study cohort as a random intercept.

[d] *P* values from Wald tests that compared fixed effect models with and without study by metabolite interaction term.

eFigure 4. Comparison of main effects (odds ratios) when excluding individual studies



Odds ratios and 95% confidence intervals represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual phthalate metabolite. Associations were estimated by multiple logistic regression models that included all participants (Overall), or excluded participants from each study. All models adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI. Studies are ordered by the relative size of the study population

eTable 9. Effect estimates and Wald tests for tests of heterogeneity in confounding by study

			Heterogeneity in confounding *P* values from Wald tests of interaction models [b]							
		Primary model	Study*Age		Study*prepregnancy BMI		Study*Race/Ethnicity[c]		Study*Education[c]	
Metabolite	n	OR (95% CI)	OR (95% CI)	Wald	OR (95% CI)	Wald	OR (95% CI)	Wald	OR (95% CI)	Wald
MEP	6045	1.07 (0.93,1.24)	1.08 (0.94,1.24)	0.42	1.07 (0.93,1.23)	0.98	1.26 (1.02,1.56)	0.92	1.07 (0.92,1.26)	0.50
MBP	6045	1.12 (0.98,1.27)	1.11 (0.97,1.27)	0.44	1.12 (0.98,1.28)	0.98	1.03 (0.84,1.27)	0.93	1.09 (0.92,1.29)	0.56
MiBP	5991	1.16 (1.00,1.34)	1.16 (1.00,1.35)	0.40	1.16 (1.00,1.35)	0.96	1.02 (0.82,1.28)	0.94	1.15 (0.96,1.38)	0.56
MBzP	5991	0.98 (0.83,1.14)	0.98 (0.83,1.15)	0.42	0.98 (0.83,1.14)	0.97	0.95 (0.75,1.19)	0.94	0.98 (0.81,1.18)	0.55
MEHP	6045	1.04 (0.91,1.19)	1.04 (0.90,1.19)	0.42	1.04 (0.91,1.20)	0.98	1.03 (0.86,1.23)	0.94	1.00 (0.85,1.18)	0.48
MEHHP	6045	1.03 (0.90,1.19)	1.04 (0.90,1.19)	0.42	1.04 (0.91,1.19)	0.98	0.94 (0.77,1.16)	0.94	0.91 (0.76,1.10)	0.39
MECPP	5471	1.16 (1.00,1.34)	1.16 (1.00,1.34)	0.36	1.17 (1.01,1.35)	0.95	1.26 (1.04,1.53)	0.93	1.19 (1.00,1.43)	0.49
MEOHP	6045	1.00 (0.88,1.15)	1.00 (0.88,1.15)	0.42	1.01 (0.88,1.15)	0.98	0.99 (0.81,1.20)	0.94	0.95 (0.81,1.12)	0.48
MCPP	5673	1.14 (1.01,1.29)	1.14 (1.01,1.28)	0.36	1.14 (1.01,1.29)	0.96	1.14 (0.97,1.34)	0.94	1.14 (0.99,1.31)	0.39
MCOP	3758	1.04 (0.84,1.29)	1.04 (0.84,1.29)	0.47	1.04 (0.84,1.29)	0.97	1.09 (0.80,1.50)	0.99	1.10 (0.86,1.42)	0.70
MCNP	3794	1.06 (0.92,1.24)	1.06 (0.91,1.23)	0.45	1.07 (0.92,1.24)	0.98	1.12 (0.91,1.36)	1.00	1.12 (0.94,1.35)	0.71

[a] OR and 95% confidence interval (CI) represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual biomarker. Associations were estimated by multiple logistic regression models. All models adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI.

[b] *P* values from Wald tests that compared models with and without designated interaction term.

[c] Models testing interactions between study and a categorical confounder (i.e., maternal race/ethnicity and education) required fitting a different subset of participants due to small subcategory sample sizes within individual studies. Thus, studies with limited to no confounder strata variation (e.g., race/ethnicity among PROTECT) were dropped from certain metabolite-specific models.

eTable 10. Comparison of odds ratio (OR) estimates for preterm birth with additional adjustment for year of delivery, maternal smoking, and parity

Metabolite	n	Primary model OR (95% CI)[a]	Delivery Year OR (95% CI)[a,b]	Smoking OR (95% CI)[a,b]	Parity OR (95% CI)[a,b]
MEP	6045	1.07 (0.93, 1.24)	1.07 (0.93, 1.23)	1.07 (0.93, 1.24)	1.08 (0.93, 1.24)
MBP	6045	1.12 (0.98, 1.27)	1.11 (0.97, 1.27)	1.12 (0.98, 1.27)	1.12 (0.98, 1.27)
MiBP	5991	1.16 (1.00, 1.34)	1.17 (1.00, 1.36)	1.16 (1.00, 1.34)	1.16 (1.00, 1.34)
MBzP	5991	0.98 (0.83, 1.14)	0.97 (0.83, 1.13)	0.98 (0.83, 1.14)	0.97 (0.83, 1.14)
MEHP	6045	1.04 (0.91, 1.19)	1.03 (0.89, 1.17)	1.04 (0.91, 1.19)	1.04 (0.91, 1.19)
MEHHP	6045	1.03 (0.90, 1.19)	1.02 (0.89, 1.17)	1.03 (0.90, 1.19)	1.03 (0.90, 1.19)
MECPP	5471	1.16 (1.00, 1.34)	1.14 (0.98, 1.33)	1.16 (1.00, 1.34)	1.16 (1.00, 1.34)
MEOHP	6045	1.00 (0.88, 1.15)	0.99 (0.86, 1.13)	1.00 (0.88, 1.15)	1.00 (0.88, 1.15)
MCPP	5673	1.14 (1.01, 1.29)	1.14 (1.00, 1.28)	1.14 (1.01, 1.29)	1.14 (1.01, 1.29)
MCOP	3758	1.04 (0.84, 1.29)	1.06 (0.85, 1.31)	1.04 (0.84, 1.29)	1.04 (0.84, 1.29)
MCNP	3794	1.06 (0.92, 1.24)	1.06 (0.91, 1.23)	1.06 (0.92, 1.24)	1.06 (0.92, 1.24)

[a] OR and 95% confidence interval (CI) represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual biomarker. Associations were estimated by multiple logistic regression models.

[a] Primary model adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI.

[b] Same adjustment as primary model[a], but additionally adjusted for the respective variable listed, including: categorical variable based on year of delivery (i.e., 1983-2000, 2001-2010, or 2011-2018); dichotomous variable based on any level of maternal smoking in pregnancy (i.e., yes or no); or dichotomous variable for parity (i.e., nulliparous or parous).

eTable 11. Odds ratio (OR) for preterm birth in the overall study population and stratified by fetal sex

Metabolite	n[a]	OR (95%CI)[b]	Wald[c]
MEP			
Overall	6045	1.07 (0.93, 1.23)	0.72
Female	2899	1.08 (0.87, 1.33)	
Male	3146	1.06 (0.88, 1.29)	
MBP			
Overall	6045	1.12 (0.98, 1.28)	0.33
Female	2899	1.22 (1.00, 1.50)	
Male	3146	1.04 (0.87, 1.24)	
MiBP			
Overall	5991	1.16 (1.00, 1.35)	0.21
Female	2873	1.13 (0.90, 1.42)	
Male	3118	1.17 (0.96, 1.44)	
MBzP			
Overall	5991	0.98 (0.83, 1.14)	0.37
Female	2873	1.06 (0.83, 1.34)	
Male	3118	0.90 (0.73, 1.12)	
MEHP			
Overall	6045	1.04 (0.90, 1.19)	0.71
Female	2899	1.01 (0.83, 1.23)	
Male	3146	1.05 (0.87, 1.26)	
MEHHP			
Overall	6045	1.04 (0.90, 1.19)	0.72
Female	2899	1.05 (0.85, 1.29)	
Male	3146	1.02 (0.85, 1.23)	
MECPP			
Overall	5471	1.16 (1.00, 1.34)	0.85
Female	2640	1.16 (0.94, 1.45)	
Male	2831	1.17 (0.96, 1.44)	
MEOHP			
Overall	6045	1.01 (0.88, 1.15)	0.97
Female	2899	0.99 (0.81, 1.20)	
Male	3146	1.02 (0.85, 1.22)	
MCPP			
Overall	5673	1.14 (1.01, 1.29)	0.24
Female	2740	1.18 (0.99, 1.42)	
Male	2933	1.11 (0.94, 1.31)	
MCOP			
Overall	3758	1.05 (0.84, 1.30)	0.44
Female	1801	0.95 (0.69, 1.32)	
Male	1957	1.15 (0.86, 1.54)	
MCNP			
Overall	3794	1.07 (0.92, 1.24)	0.96
Female	1817	1.06 (0.85, 1.32)	
Male	1977	1.10 (0.88, 1.36)	

[a] Stratum-specific sample size (n) varied between imputations.

[b] OR and 95% confidence interval (CI) represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual biomarker. Associations were estimated by multiple logistic regression models. All models adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI. The overall model additionally adjusted for maternal fetal sex (male/female) to allow for Wald test estimates of nested models.

[c] *P* values from Wald tests come from tests of nested models that included an interaction between phthalate biomarker and fetal sex.

eTable 12. Urinary phthalate metabolite specified using non-linear term

		Quadratic term[a,b]	
Metabolite	n	OR (95% CI)[a]	*P* value
MEP	6045	0.99 (0.88, 1.10)	0.81
MBP	6045	1.06 (0.97, 1.16)	0.22
MiBP	5991	1.02 (0.97, 1.08)	0.41
MBzP	5991	0.93 (0.82, 1.06)	0.29
MEHP	6045	1.02 (0.94, 1.12)	0.63
MEHHP	6045	1.03 (0.94, 1.13)	0.49
MECPP	5471	0.95 (0.87, 1.04)	0.26
MEOHP	6045	1.00 (0.92, 1.08)	0.91
MCPP	5673	1.02 (0.95, 1.09)	0.65
MCOP	3758	0.95 (0.78, 1.16)	0.63
MCNP	3794	1.05 (0.98, 1.13)	0.15

[a] OR and 95% confidence interval (CI) represent estimated odds of preterm birth compared to term birth per interquartile range increase in individual biomarker. Associations were estimated by multiple logistic regression models. Primary model adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI.

[b] Metabolite concentrations were specified as linear and quadratic terms. The coefficient and *P* value for the quadratic term are shown.

eTable 13. Estimated change (β) in length of gestation (weeks) per IQR increase in urinary phthalate biomarkers

Metabolite	n	Change in length of gestation (weeks, 95% CI)[a]
MEP	6045	-0.03 (-0.10,0.04)
MBP	6045	-0.09 (-0.16,-0.03)
MiBP	5991	-0.08 (-0.15,-0.01)
MBzP	5991	-0.07 (-0.14,0.00)
MEHP	6045	-0.01 (-0.07,0.05)
MEHHP	6045	-0.03 (-0.10,0.03)
MECPP	5471	-0.06 (-0.13,0.01)
MEOHP	6045	-0.01 (-0.07,0.06)
MCPP	5673	-0.05 (-0.10,0.01)
MCOP	3758	-0.05 (-0.14,0.04)
MCNP	3794	-0.01 (-0.07,0.06)

[a] Multiple linear regression models specified study cohort as categorical covariate. Sampling weights were implemented to account for LIFECODES case-control study design. All models adjusted for maternal age, race/ethnicity, education, and prepregnancy BMI.

eReferences

1. Yazdy MM, Coull BA, Gardiner JC, et al. A possible approach to improving the reproducibility of urinary concentrations of phthalate metabolites and phenols during pregnancy. *J Expo Sci Environ Epidemiol*. Sep 2018;28(5):448-460. doi:10.1038/s41370-018-0050-0

2. Buckley JP, Palmieri RT, Matuszewski JM, et al. Consumer product exposures associated with urinary phthalate levels in pregnant women. *J Expo Sci Environ Epidemiol*. Sep 2012;22(5):468-75. doi:10.1038/jes.2012.33

3. Barrett ES, Velez M, Qiu X, Chen SR. Reducing Prenatal Phthalate Exposure Through Maternal Dietary Changes: Results from a Pilot Study. *Matern Child Health J*. Sep 2015;19(9):1936-42. doi:10.1007/s10995-015-1707-0

4. Martina CA, Weiss B, Swan SH. Lifestyle behaviors associated with exposures to endocrine disruptors. *Neurotoxicology*. Dec 2012;33(6):1427-1433. doi:10.1016/j.neuro.2012.05.016

5. Robledo CA, Peck JD, Stoner J, et al. Urinary phthalate metabolite concentrations and blood glucose levels during pregnancy. *Int J Hyg Environ Health*. May 2015;218(3):324-30. doi:10.1016/j.ijheh.2015.01.005

6. LaRocca J, Binder AM, McElrath TF, Michels KB. The impact of first trimester phthalate and phenol exposure on IGF2/H19 genomic imprinting and birth outcomes. *Environ Res*. Aug 2014;133:396-406. doi:10.1016/j.envres.2014.04.032

7. Adibi JJ, Hauser R, Williams PL, et al. Maternal urinary metabolites of Di-(2-Ethylhexyl) phthalate in relation to the timing of labor in a US multicenter pregnancy cohort study. *Am J Epidemiol*. Apr 15 2009;169(8):1015-24. doi:10.1093/aje/kwp001

8. Weinberger B, Vetrano AM, Archer FE, et al. Effects of maternal exposure to phthalates and bisphenol A during pregnancy on gestational age. *J Matern Fetal Neonatal Med*. Mar 2014;27(4):323-7. doi:10.3109/14767058.2013.815718

9. Wolff MS, Engel SM, Berkowitz GS, et al. Prenatal phenol and phthalate exposures and birth outcomes. *Environ Health Perspect*. Aug 2008;116(8):1092-7. doi:10.1289/ehp.11007

10. Whyatt RM, Adibi JJ, Calafat AM, et al. Prenatal di(2-ethylhexyl)phthalate exposure and length of gestation among an inner-city cohort. *Pediatrics*. Dec 2009;124(6):e1213-20. doi:10.1542/peds.2009-0325

11. Shoaff JR, Romano ME, Yolton K, Lanphear BP, Calafat AM, Braun JM. Prenatal phthalate exposure and infant size at birth and gestational duration. *Environ Res*. Oct 2016;150:52-58. doi:10.1016/j.envres.2016.05.033

12. van Buuren S. *Flexible Imputation of Missing Data. Second Edition*. Chapman & Hall/CRC; 2018. https://stefvanbuuren.name/fimd/index.html

13. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw*. Dec 2011;45(3):1-67.

14. O'Brien KM, Upson K, Cook NR, Weinberg CR. Environmental Chemicals in Urine and Blood: Improving Methods for Creatinine and Lipid Adjustment. *Environ Health Perspect*. Feb 2016;124(2):220-7. doi:10.1289/ehp.1509693

15. Kuiper JR, O'Brien KM, Ferguson KK, Buckley JP. Urinary specific gravity measures in the U.S. population: Implications for the adjustment of non-persistent chemical urinary biomarker data. *Environ Int*. May 29 2021;156:106656. doi:10.1016/j.envint.2021.106656

16. Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. *Environ Health Perspect*. Feb 2005;113(2):192-200. doi:10.1289/ehp.7337

17. MacPherson S, Arbuckle TE, Fisher M. Adjusting urinary chemical biomarkers for hydration status during pregnancy. *J Expo Sci Environ Epidemiol*. Sep 2018;28(5):481-493. doi:10.1038/s41370-018-0043-z

18. Yeh HC, Lin YS, Kuo CC, et al. Urine osmolality in the US population: implications for environmental biomonitoring. *Environ Res*. Jan 2015;136:482-90. doi:10.1016/j.envres.2014.09.009

19. Berger K, Eskenazi B, Balmes J, et al. Prenatal high molecular weight phthalates and bisphenol A, and childhood respiratory and allergic outcomes. *Pediatr Allergy Immunol*. Feb 2019;30(1):36-46. doi:10.1111/pai.12992

20. Keil AP, Buckley JP, Kalkbrenner AE. Bayesian G-Computation for Estimating Impacts of Interventions on Exposure Mixtures: Demonstration With Metals From Coal-Fired Power Plants and Birth Weight. *Am J Epidemiol*. Dec 1 2021;190(12):2647-2657. doi:10.1093/aje/kwab053

21. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. Dec 2009;38(6):1599-611. doi:10.1093/ije/dyp192

22. Garcia E, Urman R, Berhane K, McConnell R, Gilliland F. Effects of policy-driven hypothetical air pollutant interventions on childhood asthma incidence in southern California. *Proc Natl Acad Sci U S A*. Aug 6 2019;116(32):15883-15888. doi:10.1073/pnas.1815678116

23. Gennings C, Svensson K, Wolk A, Lindh C, Kiviranta H, Bornehag CG. Using Metrics of a Mixture Effect and Nutrition from an Observational Study for Consideration towards Causal Inference. *Int J Environ Res Public Health*. Feb 17 2022;19(4)doi:10.3390/ijerph19042273

24. Westreich D, Edwards JK, Rogawski ET, Hudgens MG, Stuart EA, Cole SR. Causal Impact: Epidemiological Approaches for a Public Health of Consequence. *Am J Public Health*. Jun 2016;106(6):1011-2. doi:10.2105/AJPH.2016.303226

25. Raz R, Kioumourtzoglou MA, Weisskopf MG. Live-Birth Bias and Observed Associations Between Air Pollution and Autism. *Am J Epidemiol*. Nov 1 2018;187(11):2292-2296. doi:10.1093/aje/kwy172

26. Ferguson KK, Lan Z, Yu Y, Mukherjee B, McElrath TF, Meeker JD. Urinary concentrations of phenols in association with biomarkers of oxidative stress in pregnancy: Assessment of effects independent of phthalates. *Environ Int*. Oct 2019;131:104903. doi:10.1016/j.envint.2019.104903

27. Chia AR, Chen LW, Lai JS, et al. Maternal Dietary Patterns and Birth Outcomes: A Systematic Review and Meta-Analysis. *Adv Nutr*. Jul 1 2019;10(4):685-695. doi:10.1093/advances/nmy123

28. Sathyanarayana S, Alcedo G, Saelens BE, et al. Unexpected results in a randomized dietary trial to reduce phthalate and bisphenol A exposures. *J Expo Sci Environ Epidemiol*. Jul 2013;23(4):378-84. doi:10.1038/jes.2013.9

29. Silva MJ, Samandar E, Preau JL, Jr., Reidy JA, Needham LL, Calafat AM. Quantification of 22 phthalate metabolites in human urine. *J Chromatogr B Analyt Technol Biomed Life Sci*. Dec 1 2007;860(1):106-12. doi:10.1016/j.jchromb.2007.10.023

30. Calafat AM *Phthalate Metabolites Method 6306.03*. Centers for Disease Control and Prevention. Atlanta, GA. 2010.

31. Kato K, Silva MJ, Needham LL, Calafat AM. Determination of 16 phthalate metabolites in urine using automated sample preparation and on-line preconcentration/high-performance liquid chromatography/tandem mass spectrometry. *Anal Chem*. May 1 2005;77(9):2985-91. doi:10.1021/ac0481248

32. Silva MJ, Slakman AR, Reidy JA, et al. Analysis of human urine for fifteen phthalate metabolites using automated solid-phase extraction. *J Chromatogr B Analyt Technol Biomed Life Sci*. Jun 5 2004;805(1):161-7. doi:10.1016/j.jchromb.2004.02.038

33. Lewis RC, Meeker JD, Peterson KE, et al. Predictors of urinary bisphenol A and phthalate metabolite concentrations in Mexican children. *Chemosphere*. Nov 2013;93(10):2390-8. doi:10.1016/j.chemosphere.2013.08.038