United States
Environmental Protection Agency

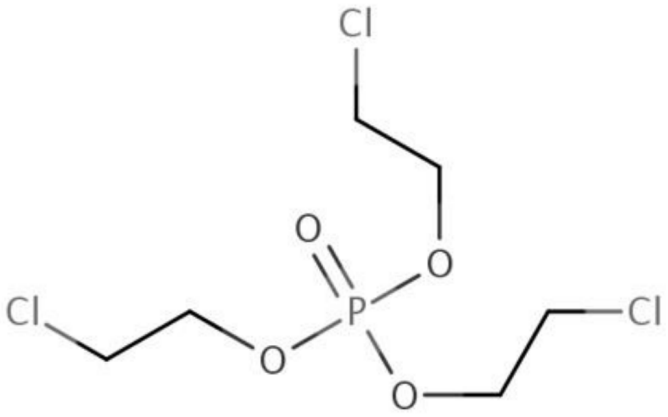
September 2024
Office of Chemical Safety and
Pollution Prevention

Risk Evaluation for Tris(2-chloroethyl) Phosphate (TCEP)

Systematic Review Supplemental File:

Data Quality Evaluation Information for Human Health Hazard Epidemiology

CASRN: 115-96-8



September 2024

This supplemental file contains the data quality evaluation results for epidemiology data sources that met the PECO screening criteria in the *Risk Evaluation for Tris(2-chloroethyl) phosphate (TCEP)*. EPA conducted data quality evaluation based on author-reported descriptions and results; additional analyses (*e.g.*, statistical analyses performed during data integration into the risk evaluation) potentially conducted by EPA are not contained in this supplemental file. EPA used the TSCA systematic review process described in the *Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances*(also referred to as '2021 Draft Systematic Review Protocol'). Any updated steps in the systematic review process since the publication of the 2021 Draft Systematic Review Protocol are described in *Risk Evaluation for Tris(2-chloroethyl) phosphate (TCEP) – Systematic Review Protocol*.

Table of Contents

Metabolite: Bis-2-chloroethyl phosphate (BCEP)

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-wheeze
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	The study was conducted among elementary school students in Sapporo, Japan in 2008. The survey was sent to 6393 school children across 12 public elementary schools, and 44008 students responded, with 951 students interested in participating. Only 681 families were still at the same elementary school in 2009, of those only 128 families were able to be contacted for a home visit, meaning there was an overall participation rate of 2.9%Despite a low participation rate, the reasons for exclusion are clearly defined at each step and are unlikely to introduce significant bias.
	Metric 2:	Attrition	High	Of the final 128 participants, none were excluded from the data analysis after home visits/data collection occurred. There is no missing exposure or outcome data.
	Metric 3:	Comparison Group	Medium	Participants in this cross-sectional study appear to be similar in terms of baseline characteristics. Most variables are controlled for in statistical analyses. Height and weight are not controlled for, but the listed mean + standard deviations imply that there is a somewhat large range.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure is reported as TCEP measured from urine.
	Metric 5:	Exposure Levels	Medium	The study reports on a range of exposures from the LOQ to 1.13 nM of TCEP. In statistical analyses, TCEP concentrations are split into tertiles.
	Metric 6:	Temporality	Low	Exposure was measured after the onset of symptoms, so the temporality of the exposure and outcome is uncertain. However, due to the ubiquity of TCEP in the environment, it may be reasonable to assume that they were exposed to the same amount of TCEP prior to exposure assessment. However, the study does not check to see if children had moved recently, which could have altered exposure levels.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes were assessed via the International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire, which was filled out by study investigators and the parents of children. The questionnaire asks for symptoms common for wheeze instead of actual medical diagnoses. Parents may also be impacted by the desirability bias. However, this is likely to have a non-differential effect.
	Metric 8:	Reporting Bias	High	A description of the the outcomes is clearly mentioned in the methods. Results are presented as odds ratios with 95% confidence intervals and p-values.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-wheeze
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	The results are reported as odds ratios that are adjusted for sex, grade, annual income, and the dampness index of the child's home. They were all included based on a priori evidence.
	Metric 10:	Covariate Characterization	Medium	Information on confounders were obtained via questionnaire. Sex and grade were measured via questionnaire and were unlikely to be reported incorrectly. Annual income is also self-reported, and could fall victim to desirability bias. There were also missing values for income (14.8%) to which the mean annual household income was assigned. Finally, a "dampness" index was calculated by study investigator observation of dampness-related problems in each dwelling, such as condensation and visible mold.
	Metric 11:	Co-exposure Counfounding	Medium	The intent of this study was to measure co-exposures of phthalates and phosphate flame retardants and allergic outcomes. Statistical models were used in this study to examine in urine individual phthalates, combinations of phthalates, metabolites of PFRs and a combination of certain PFRs.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The use of a cross-sectional design to understand the relationship between TCEP and wheeze/allergy symptoms is an appropriate study design, and calculating odds ratios via logistic regression is an appropriate method.
	Metric 13:	Statistical Power	Medium	The number of participants in each tertile is >20, which could be sufficiently large to detect an effect. The authors do not calculate statistical power, however they do mention that the power may not be enough to detect a significant effect. They may be referring to the multipollutant models in this case.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and allows for replication given the study data.
	Metric 15:	Statistical Analysis	High	The model building is transparent and it is clear why variables were chosen.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	TCEP is a parent compound.
	Metric 17:	Effect Biomarker	N/A	Not applicable - no biomarker of effect.
	Metric 18:	Method Sensitivity	Medium	The % of samples below the detection limit is stated to be 14.8%, which is not incredibly high but is sufficient to address the research hypothesis.
	Metric 19:	Biomarker Stability	Medium	There is no documented stability data, but is clarified that spot urine samples were collected in polypropylene containers and refrigerated until the study visit, and then stored a -20 deg C until the day of analysis.
	Metric 20:	Sample Contamination	Medium	There is no documentation of potential sample contamination.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-wheeze
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	TCEP was measured using liquid chromatography with tandem mass spectrometry (LC-MS/MS).
	Metric 22:	Matrix Adjustment	Medium	TCEP concentrations were creatinine-adjusted, and ranges are reported without adjustment and with adjustment. However, the statistical model only uses adjusted concentrations of TCEP.

Additional Comments: This cross-sectional study measures the association between TCEP measured in spot urine samples and parent-reported symptoms of wheeze. Associations were analyzed using logistic regression and the study calculated odds ratios. No significant associations in single-pollutant models for TCEP were reported. The study has some deficiencies in outcome assessment, as there is no verifiable method used in this study.

Overall Quality Determination High

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-eczema, allergic rhinoconjunctivitis
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	The study was conducted among elementary school students in Sapporo, Japan in 2008. The survey was sent to 6393 school children across 12 public elementary schools, and 44008 students responded, with 951 students interested in participating. Only 681 families were still at the same elementary school in 2009, of those only 128 families were able to be contacted for a home visit, meaning there was an overall participation rate of 2.9%Despite a low participation rate, the reasons for exclusion are clearly defined at each step and are unlikely to introduce significant bias.
	Metric 2:	Attrition	High	Of the final 128 participants, none were excluded from the data analysis after home visits/data collection occurred. There is no missing exposure or outcome data.
	Metric 3:	Comparison Group	Medium	Participants in this cross-sectional study appear to be similar in terms of baseline characteristics. Most variables are controlled for in statistical analyses. Height and weight are not controlled for, but the listed mean + standard deviations imply that there is a somewhat large range.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure is reported as TCEP measured from urine.
	Metric 5:	Exposure Levels	Medium	The study reports on a range of exposures from the LOQ to 1.13 nM of TCEP. In statistical analyses, TCEP concentrations are split into tertiles.
	Metric 6:	Temporality	Low	Exposure was measured after the onset of symptoms, so the temporality of the exposure and outcome is uncertain. However, due to the ubiquity of TCEP in the environment, it may be reasonable to assume that they were exposed to the same amount of TCEP prior to exposure assessment. However, the study does not check to see if children had moved recently, which could have altered exposure levels.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes were assessed via the International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire, which was filled out by the parents. The questionnaire asks for symptoms common for wheeze, eczema, and allergic rhinoconjunctivitis. There was no medical diagnosis by a physician.
	Metric 8:	Reporting Bias	High	A description of the the outcomes is clearly mentioned in the methods. Results are presented as odds ratios with 95% confidence intervals and p-values.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-eczema, allergic rhinoconjunctivitis
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	The results are reported as odds ratios that are adjusted for sex, grade, annual income, and the dampness index of the child's home. They were all included based on a priori evidence.
	Metric 10:	Covariate Characterization	Medium	Information on confounders were obtained via questionnaire. Sex and grade were measured via questionnaire and were unlikely to be reported incorrectly. Annual income is also self-reported, and could fall victim to desirability bias. There were also missing values for income (14.8%) to which the mean annual household income was assigned. Finally, a "dampness" index was calculated by study investigator observation of dampness-related problems in each dwelling, such as condensation and visible mold.
	Metric 11:	Co-exposure Counfounding	Medium	The intent of this study was to measure co-exposures of phthalates and phosphate flame retardants and allergic outcomes. Statistical models were used in this study to examine in urine individual phthalates, combinations of phthalates, metabolites of PFRs and a combination of certain PFRs.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The use of a cross-sectional design to understand the relationship between TCEP and wheeze/allergy symptoms is an appropriate study design, and calculating odds ratios via logistic regression is an appropriate method.
	Metric 13:	Statistical Power	Medium	The number of participants in each tertile is >20, which could be sufficiently large to detect an effect. The authors do not calculate statistical power, however they do mention that the power may not be enough to detect a significant effect. They may be referring to the multipollutant models in this case.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and allows for replication given the study data.
	Metric 15:	Statistical Analysis	High	The model building is transparent and it is clear why variables were chosen.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	TCEP is a parent compound.
	Metric 17:	Effect Biomarker	N/A	Not applicable - no biomarker of effect.
	Metric 18:	Method Sensitivity	Medium	The % of samples below the detection limit is stated to be 14.8%, which is not incredibly high but is sufficient to address the research hypothesis.
	Metric 19:	Biomarker Stability	Medium	There is no documented stability data, but is clarified that spot urine samples were collected in polypropylene containers and refrigerated until the study visit, and then stored a -20 deg C until the day of analysis.
	Metric 20:	Sample Contamination	Medium	There is no documentation of potential sample contamination.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Araki, A., Bamai, Y. A., Bastiaensen, M., Eede, V.d., N., Kawai, T., Tsuboi, T., Miyashita, C., Itoh, S., Goudarzi, H., Konno, S., Covaci, A., Kishi, R. (2020). Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. Environmental Research 183:109212.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-eczema, allergic rhinoconjunctivitis
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6957526

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	TCEP was measured using liquid chromatography with tandem mass spectrometry (LC-MS/MS).
	Metric 22:	Matrix Adjustment	Medium	TCEP concentrations were creatinine-adjusted, and ranges are reported without adjustment and with adjustment.However, the statistical model only uses adjusted concentrations of TCEP.

Additional Comments: This cross-sectional study measures the association between TCEP measured in spot urine samples and parent-reported symptoms of allergic rhinoconjunctivitis and eczema. Associations were analyzed using logistic regression and the study calculated odds ratios. No significant associations in single-pollutant models for TCEP were reported. The study has some deficiencies in outcome assessment, as there is no verifiable method used in this study.

Overall Quality Determination High

Study Citation:	Canbaz, D., Velzen, van, M. J., Hallner, E., Zwinderman, A. H., Wickman, M., Leonards, P. E., Ree, van, R., Rijt, van, L. S. (2015). Exposure to organophosphate and polybrominated diphenyl ether flame retardants via indoor dust and childhood asthma. Indoor Air 26(3):403-413.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Asthma
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	2994738

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	The authors reported participant selection in both this study and in Almqvist et al., 2003. However, not all of the elements were reported, such as participation rate at all phases of the study.
	Metric 2:	Attrition	Medium	Exclusion and missing values were reported in the Materials and Methods Section, Table 1, and Table S1.
	Metric 3:	Comparison Group	High	Study matches controls based on sex, atopic background of the parents, and socioeconomic status. Study states that the asthmatic children and their matched controls did not differ significantly according to several sociodemographic characteristics outlined in Table S1.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Samples were filtered, sealed, and stored appropriately. Analyzed by GC-EI-MS with QA detailed in Brandsma et al., 2014. Additional details in Almqvist et al., 2003
	Metric 5:	Exposure Levels	Medium	Table 3 provides the range and distribution of exposure for both cases and controls.
	Metric 6:	Temporality	High	House dust adequate to capture short half-life of TCEP. Dust was collected and analyzed when children were two months of age; exposure precedes disease.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Asthma defined according to set of criteria, including doctor's diagnosis and asthma medicine prescription.
	Metric 8:	Reporting Bias	Low	Not all data shown from analyses, including the multivariate linear regression analysis.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study reports multivariate linear regression analyses to adjust for covariates; however, the data was not shown in the main paper or supplemental. Study noted that the results did not differ from those reported.
	Metric 10:	Covariate Characterization	High	Questionnaire and doctors' diagnoses used.
	Metric 11:	Co-exposure Counfounding	Medium	Adjusted for in the multivariate linear regression analyses; data not shown but stated to not affect results.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Case-control used but no logistic regression conducted; no odds ratios.
	Metric 13:	Statistical Power	Medium	110 cases; 110 controls; adequate sample size

Continued on next page . . .

. . . continued from previous page

Study Citation:	Canbaz, D., Velzen, van, M. J., Hallner, E., Zwinderman, A. H., Wickman, M., Leonards, P. E., Ree, van, R., Rijt, van, L. S. (2015). Exposure to organophosphate and polybrominated diphenyl ether flame retardants via indoor dust and childhood asthma. Indoor Air 26(3):403-413.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Asthma
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	2994738

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	Sufficient summary of analyses
	Metric 15:	Statistical Analysis	High	Methods are transparent

Additional Comments: None

Overall Quality Determination Medium

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	This study analyzed associations between TCEP in household dust collected shortly after birth and maternal mental health during pregnancy and postpartum using a sub-sample of from the CHILD (Canadian Healthy Infant Longitudinal Development) study.The ”CHILD Cohort Study recruited adult pregnant women who were in their second trimester from health centers located across Canada.” Eligibility criteria included being able to communicate in English and the infant being born at 35 weeks gestation or later.The analysis sample was compiled for a nested case-cohort study that initially focused on exposure to phthalates and respiratory outcomes in children. Of the 3,628 CHILD participants, 2,319 were potentially eligible for inclusion based on having house dust samples, along with having genotyping from a related study and the child having been clinically assessed at age 5 years. Out of 3628 participants, 2319 met these criteria, and a random sample of 436 of these 2319 eligible children were selected for inclusion. Additionally, all remaining children with asthma or recurrent wheeze (n=290), were included. Eight otherwise eligible participants were excluded due to issues analyzing their house dust samples. The final sample comprised 718 participants.Distributions of characteristics are presented in Supplemental Table S1; The authors stated “characteristics of this sub-sample were similar to those in the full cohort, although missingness in the sub-sample was generally lower than in the full cohort”; statistical significance testing of differences was not discussed. There was no direct evidence of selection bias. However, a potentially important concern is that selectivity related to oversampling for childhood respiratory illnesses was not addressed in the analyses, e.g., by using selection probability weights or analyzing associations separately in the random sample. A publication from this cohort found house dust TCEP to be non-linearly associated with lower odds of asthma and recurrent wheeze at age 5 years (significant for Q3 vs Q1; Navaranjan et al, 2021 HEROID 10134087), suggesting that selection based on child respiratory outcomes may not be independent of TCEP. Childhood asthma has been associated with maternal asthma and maternal postpartum depression (e.g., Blais et al. 2019 PMID 30292921; Alcala et al 2023 PMID 36128727). Given these relationships, potential bias (due to collider stratification) is a concern: the TCEP-maternal depression association may differ among children with respiratory outcomes vs the general population. However, as there was no direct evidence that oversampling for childhood respiratory illness affected the validity or significance of associations between house dust TCEP and maternal mental health outcomes, a medium rating was selected.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
	Metric 2:	Attrition	Low	The study analyzed maternal mental health outcomes measured at four time points, over which the proportion of participants with available data varied (57.1% overall, and 77.4%, 67.8%, 63.6%, and 70.2% at 18 weeks' gestation, 36 weeks' gestation, 6 months postpartum, and one year postpartum, respectively). A potentially important concern is that over time, the proportion of women who did not report mental health scores increased. The proportion of non-reporters at 18 weeks gestation vs. 6 months postpartum increased from 6.3 % to 17.3% for depression scores and from 6.3% to 17.0% for stress scores. This >10% increase in missing outcome reporting is similar in magnitude to the prevalence of clinically relevant depression (12.4% at 6 months postpartum) and several times higher than the prevalence of elevated stress (2.2% at 6 months). The authors made an effort to address missingness using multiple imputation by chained equations. However, imputation models used fixed characteristics (e.g. age, education) which may not adequately predict the significant within-person changes in mental health scores across time points. Potential bias as a consequence of the >10% increase over time in outcome non-reporting is an important concern.
	Metric 3:	Comparison Group	Medium	The analysis compared mental health scores among women from households with higher vs. lower levels of TCEP in house dust. Although potential bias due to over-sampling for childhood respiratory illnesses cannot be excluded as noted under Metric 1, the nested prospective cohort design minimizes potential differences between groups, and differences between groups were considered as potential confounding variables. There was no evidence for additional concerns related to the comparison group.

Domain 2: Exposure Characterization

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Low	TCEP and several other organophosphate ester (OPE) flame retardants were measured in finely sieved house dust using gas chromatography. Samples were collected at a single time point (3-4 months postpartum) by vacuuming using a standardized protocol. Laboratory analyses were based on homogenized equal aliquots combined from two areas: (i) the floor of the most used living area; and (ii) the child's mattress sheet or cover plus either an adjacent 2-meter area of carpeting or the whole floor if uncarpeted. Only 1.4% of samples had TCEP concentrations below the LOD; values below the LOD (1.4%) were imputed using the NDExpo web application. Samples were collected from each area using new thimbles that were transported in sterile glass bottles (Navaranjan et al, 2021 HEROID 10134087). Samples were stored at room temperature for two weeks, then at -80 degrees C prior to analysis. There was no evidence to suggest concerns related to sample degradation. Dust sample collection was standardized, and standard amounts of homogenized aliquots were analyzed. The analyses used gas chromatography [Agilent GC-MSD (6890-5975)] operating in electron impact mode using selective ion monitoring. TCEP exposure was analyzed as concentrations (ng/g) in dust. Quality control procedures included using standard reference material and analyzing control samples and laboratory blanks. Blanks with OPE concentrations below detection limits were appropriately included. A mixture of isotopically labelled OPEs, including TCEP-d12, were added as surrogate standards to selected blanks (81-112% recovery). Z-scores were calculated and analyses were conducted for TCEP exposure individually, and for summary exposure to OPEs based on the sum score index method.There were several concerns with the exposure assessment for TCEP. First, it is uncertain extent to which TCEP in the pooled dust sample with 50% from infant sleeping areas reflects maternal exposure, particularly as OPE concentrations in a subsample were considerably higher in bedrooms vs living areas (Navaranjan et al 2921, 10134087). The authors speculated that exposures of mothers and infants might be similar because "parents often share a room with their child for the first three to six months of life, and their mattress would likely also have flame retardants additives". Second, individual and household factors that may influence the extent to which these dust measures represent maternal exposure were not taken into account (e.g., cleaning frequency, frequency with which mothers slept in children's rooms, age of infant mattress vs. maternal mattress). Third, the authors acknowledge that hand wipes, wristbands, or urine samples, rather than dust samples, might have better characterized internal dose. Finally, the TCEP exposure metrics were analyzed as concentrations that did not account for variability in room size/floor area such as by calculating TCEP dust loadings, which may yield different associations (e.g., Mendy et al 2024, HEROID 11364495). Thus, there is substantial uncertainty in the extent to which the TCEP dust measures used adequately captured maternal exposures.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	TCEP exposure was analyzed using a continuous variable – concentrations per gram of sieved house dust (ng/g) converted to Z-scores. Distribution data provided suggested that the exposure variable was highly non-normal (the minimum, geometric mean, and maximum were 375, 5150 and 81200 ng/g, respectively). However, there was no direct evidence that a transformation to improve normality would have meaningfully influenced results. Thus, the range and distribution of exposure were sufficient and a continuous measure of exposure was used, which meets the requirements for a Medium rating for this metric.
	Metric 6:	Temporality	Low	Exposure was measured in house dust collected at only a single timepoint of 3-4 months postpartum. Two of the four maternal mental health assessments occurred prior to the exposure measurement, having taken place at 18- and 36 weeks' gestation. However, temporality was appropriate for the remaining two measures. Exposure was estimated using a single dust sample, and there was no information on variability in house dust TCEP over time, relative to the timing of outcome measures.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Maternal mental health outcomes were measured in pregnancy and postpartum (18 weeks' gestation, 36 weeks' gestation, 6 months postpartum, and one year postpartum). Measures were characterized using two self-administered questionnaires: the Centre for Epidemiologic Studies for Depression Scale (CES-D) was used to quantify maternal depression, and the Perceived Stress Scale (PSS) was used to quantify maternal stress. Both questionnaires have been reported to have good reliability and internal consistency in other studies in the general population. Primary analyses used continuous scores. Dichotomized scores used in supplementary analyses were informed by the literature but were not based on clinically validated cutoffs. One minor limitation in outcome characterization is that the authors did not discuss the utility and validity of these instruments to measure changes in mental health occuring specifically around the time of pregnancy. Both depression and perceived stress scores decreased from gestation to postpartum: for example, at 18 weeks' gestation vs. 6 months postpartum, mean CES-D scores were 9.3 vs. 8.3, respectively, and the prevalence of clinically relevant scores declined from 16.4% to 12.4%, respectively. The authors did not discuss the extent to which diminished scores were expected, nor whether the apparent declines might be attributable, at least in part, to increases in missing responses (from 6.3% to 17.3%, see metric 2). A related limitation is that repeated mental health scores were analyzed without characterizing whether the depression or perceived stress was persistent, of new onset, or transient. Despite these limitations, there was no evidence of important biases in outcome assessment methods that would influence the validity of the assessed associations with TCEP.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	Medium	Results were shown or described for all analyses that had been specified in the Materials and Methods section of the paper. Effect estimates were presented with confidence intervals. Numbers of participants included were provided for the primary analyses of continuous outcomes. Numbers in the dichotomous outcome groups (clinically relevant vs not clinically relevant CES-D; and Low to Moderate vs high PSS score) were shown in Table 1, but numbers of cases by TCEP exposure status were not shown for supplemental analyses of dichotomous outcomes.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Covariates were selected a priori based on the literature; use of directed acyclic graphs or other qualitative methods was not discussed. Models adjusted for study center, maternal age, household income, ethnicity, marital status, prenatal smoking, depression medication use, and season of dust collection. Maternal education and self-reported depression were considered as covariates but excluded from final models due to high collinearity; parity and delivery mode were excluded as they were less important in the literature and not significantly associated with outcomes. A potential concern is that the authors did not discuss potential confounding or effect modification by maternal physical health issues potentially related to both TCEP exposure and maternal mental health (e.g., gestational diabetes, pregnancy-induced hypertension, or maternal asthma). However, excluding these variables would be appropriate if they are intermediates. There was no evidence of important bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were measured using questionnaires, but the authors didn't specify whether the questionnaires were validated. The ambient air pollution co-exposure measures were assigned based on postal codes. Though details were limited, there was no evidence of important deficiencies or biases.
	Metric 11:	Co-exposure Counfounding	Medium	Models adjusted for annual averages for ambient air pollutants (PM2.5, NO2, and O3). A heatmap of Spearman correlations indicated that TCEP was not highly correlated with other OPEs measured in this study, suggesting that co-exposure confounding by other OPEs was unlikely. Potential confounding by other OPEs was not explicitly evaluated, although associations with exposure to multiple OPEs was analyzed using summed Z-scores for 14 individual OPEs.
Domain 5: Analysis				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric		Rating	Comments
	Metric 12:	Study Design and Methods	High	The study analyzed data nested within a prospective cohort. The primary analyses used linear mixed models for repeated continuous outcome measures and linear regression for timepoint-specific continuous outcome measures. Supplemental models of dichotomized outcomes were analyzed using mixed logistic regression and timepoint-specific logistic models. These methods are appropriate for the research question of assessing the association between exposure to TCEP and maternal depression and stress. Oversampling for childhood respiratory illness (see metric 1) could have been better addressed in the analyses, such as by using sample weights. Additional important deficiencies in design were not noted.
	Metric 13:	Statistical Power	Medium	Repeated measures analyses included 718 participants; sample sizes for timepoint specific analyses ranged from 594 to 676. There was variability in both TCEP exposure and mental health outcome variables, which were used as continuous measures in primary analyses. Although power calculations weren't presented, the number of participants appears to be adequate to detect an effect. However, relatively few women overall reported scores that were classified as elevated stress, which might have limited the power to detect differences in stress between groups when the outcome was dichotomized.
	Metric 14:	Reproducibility of Analyses	Medium	The description of analyses was adequate to be conceptually reproducible with access to the data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	High	Descriptive data were presented stratified by elevated CES-D and PSS scores. Both crude and adjusted associations were presented and the covariates for the adjusted models were specified a priori; both complete case analyses and analyses using multiple imputation were conducted. Associations between house dust TCEP concentrations and continuous mental health outcome scores were analyzed using linear mixed models for repeated outcome measures, as well as using linear regression for timepoint specific outcome measures. CES-D scores were natural log-transformed to meet the assumptions of linear regression. Dichotomized mental health outcomes were analyzed using both mixed logistic regression and timepoint-specific logistic models. Results were presented as beta coefficients or odds ratios with confidence intervals. There were several potential concerns related to data analysis, but no direct evidence of important error or bias. One potential concern is that TCEP exposure Z-scores were analyzed continuously and without transformation, assuming a linear dose-response relationship. There was no direct evidence of non-linearity in these associations. However, TCEP was non-linearly associated with childhood respiratory illnesses within this cohort (Navaranjan et al, 10134087). A second potential concern is that model diagnostics and sensitivity analyses were discussed only for analyses of summed OPE exposures. It is uncertain whether findings for TCEP were robust after exclusions such as removing participants who had moved homes during the study period. Finally, despite the aim of analyzing whether OPEs influenced depression and stress around pregnancy, sensitivity analyses for summed OPEs did not include evaluating the impact of excluding participants reporting depression prior to pregnancy.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Foster, S. A., Kile, M. L., Hystad, P., Diamond, M. L., Jantunen, L. M., Mandhane, P. J., Moraes, T. J., Navaranjan, G., Scott, J. A., Simons, E., Subbarao, P., Takaro, T. K., Turvey, S. E., Brook, J. R. (2024). Organophosphate ester flame retardants and plasticizers in house dust and mental health outcomes among Canadian mothers: A nested prospective cohort study in CHILD. Environmental Research 240(Pt 1):117451.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Maternal depression scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Centre for Epidemiologic Studies for Depression Scale (CES-D); andMaternal stress scores during pregnancy (at 18 and 36 weeks gestation) and postpartum (at 6 months and 1 year postpartum) using the Perceived Stress Scale (PSS)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11581665

Domain	Metric	Rating	Comments
Additional Comments:	This study analyzed associations between TCEP exposure, as well as summed organophosphate ester flame retardant (OPE) exposure, in household dust and maternal depression and perceived stress scores using a subset of 718 participants from the Canadian CHILD cohort. Household dust was measured at 3-4 months postpartum; mental health scores were obtained from widely-used self-administered instruments (CES-D, PSS) at weeks 18 and 36 gestation, and months 6 and 12 postpartum. Both timepoint specific and repeated measures analyses were conducted. Although most of the relevant findings were null or not statistically significant for TCEP, there was a statistically significant positive association between dust TCEP levels and odds of high perceived stress (when perceived stress was dichotomized) that was observed only at 6 months postpartum (shown in Table S10), and a small but statistically significant positive association between summed OPE levels and perceived stress in linear mixed models. However, relatively few women (n=16 to 18) reported scores that were classified as elevated stress. This study had several limitations, including the use of pooled dust samples comprised of 50% from the main living area and 50% from the infant's mattress and adjacent floor. It is uncertain to what extent these measures capture infant vs. maternal exposure. There was also an increase of more than 10% in non-reporting of mental health outcomes between pregnancy and postpartum; it is uncertain whether the apparent postpartum declines in maternal depression and perceived stress reflect improved mental health vs. selective attrition. The timing of two of the four maternal mental health measures preceded exposure sample collection. However, the authors presented timepoint specific measures to address this temporality concern. Other limitations included oversampling for childhood respiratory illnesses without incorporating sample weights despite an association between these outcomes and TCEP in the CHILD cohort, and assuming a linear dose-response relationship.		

Overall Quality Determination Medium

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participant selection was well described, including timing of study, inclusion and exclusion criteria and case ascertainment.
	Metric 2:	Attrition	High	Supplemental Figure 1 provides detailed information and shows only 1 of 71 cases excluded.
	Metric 3:	Comparison Group	High	Cases and controls were recruited from the same Health Care Center during the same time period and were matched by age and gender.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Flame retardants were measured in house dust. Collection methods and analysis of dust was fully reported. TCEP concentrations in household dust were measured by GC/EI-MS.
	Metric 5:	Exposure Levels	Low	Only 2 levels of flame retardant exposure in the home were reported for cases and controls; TCEP concentrations above the median concentration were reported for both groups. Median TCEP concentrations were shown in the box plot in Fig. 1 but concentrations were not reported. Detection limits were not reported for organophosphate flame retardants.
	Metric 6:	Temporality	Medium	Exposure at the same residence for at least 2 years prior to diagnosis; however, it is unclear whether exposure duration was sufficient for thyroid cancer.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The thyroid cancer outcome was assessed by medical review of clinical and pathology data; pathological stage was assessed based on tumor size, location, metastasis.
	Metric 8:	Reporting Bias	Medium	Health outcomes and exposures are reported, except for median exposure levels of each of the flame retardants. Adjusted ORs with CIs reported for overall incidence as well as measures of tumor aggressiveness (i.e., pathologic stage). Some data were reported for PBDEs but not the organophosphate flame retardants, such as number of detects, detection limits.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariate adjustments were described (age, income, BMI). Regression analyses were adjusted for participant age and household income. BMI was both included and excluded as a covariate. Other potential confounders were considered, but not applied because they did not alter effect estimates (race, employment status, and smoking). Ionizing radiation exposure was considered but no participants reported prior exposure.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	A questionnaire was administered to participants, but the method for assessing covariates was not described, although the study indicated that protocols were approved by an institutional review board. Did not report if the questionnaire was validated or the types of data collected in the questionnaire, but there is no evidence of confounding.
	Metric 11:	Co-exposure Counfounding	Low	Many flame retardants were measured in household dust samples. The authors acknowledge and provide data on those most highly correlated. Statistical analyses modeled each flame retardant separately. The authors indicate that component analyses were conducted to systematically assess FR mixtures but did not provide them in the article. They stated that they "did not provide any additional insights."
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The case-control study design was appropriate for the research question and applicable statistical methods were used (logistic regression models). However, the co-exposures were not accounted for in the analysis.
	Metric 13:	Statistical Power	Medium	The number of cases and controls was adequate to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis was sufficient to be reproducible with access to the analytical data.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described.

Additional Comments: For TCEP, dust levels in the home are used as surrogates for exposure, measured using MS/GC. PBDEs only were measured in serum (not organophosphate flame retardants). This case-control study used an adequate number of age- and gender-matched pairs. Multiple flame retardants were present in dust and in large ranges, depending on the flame retardant; adjustments were not made for multiple comparisons (separate models were assessed for each compound).

Overall Quality Determination High

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participant selection was well described, including timing of study, inclusion and exclusion criteria and case ascertainment.
	Metric 2:	Attrition	High	Supplemental Figure 1 provides detailed information and shows only 1 of 71 cases excluded.
	Metric 3:	Comparison Group	High	Cases and controls were recruited from the same Health Care Center during the same time period and were matched by age and gender.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Flame retardants were measured in house dust. Collection methods and analysis of dust was fully reported. TCEP concentrations in household dust were measured by GC/EI-MS.
	Metric 5:	Exposure Levels	Low	Only 2 levels of flame retardant exposure in the home were reported for cases and controls; TCEP concentrations above the median concentration were reported for both groups. Median TCEP concentrations were shown in the box plot in Fig. 1 but concentrations were not reported. Detection limits were not reported for organophosphate flame retardants.
	Metric 6:	Temporality	Medium	Exposure at the same residence for at least 2 years prior to diagnosis; however, it is unclear whether exposure duration was sufficient for thyroid cancer.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The thyroid cancer outcome was assessed by medical review of clinical and pathology data; pathological stage was assessed based on tumor size, location, metastasis.
	Metric 8:	Reporting Bias	Medium	Health outcomes and exposures are reported, except for median exposure levels of each of the flame retardants. Adjusted ORs with CIs reported for overall incidence as well as measures of tumor aggressiveness (i.e., pathologic stage). Some data were reported for PBDEs but not the organophosphate flame retardants, such as number of detects, detection limits.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariate adjustments were described (age, income, BMI). Regression analyses were adjusted for participant age and household income. BMI was both included and excluded as a covariate. Other potential confounders were considered, but not applied because they did not alter effect estimates (race, employment status, and smoking). Ionizing radiation exposure was considered but no participants reported prior exposure.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	A questionnaire was administered to participants, but the method for assessing covariates was not described, although the study indicated that protocols were approved by an institutional review board. Did not report if the questionnaire was validated or the types of data collected in the questionnaire, but there is no evidence of confounding.
	Metric 11:	Co-exposure Counfounding	Low	Many flame retardants were measured in household dust samples. The authors acknowledge and provide data on those most highly correlated. Statistical analyses modeled each flame retardant separately. The authors indicate that component analyses were conducted to systematically assess FR mixtures but did not provide them in the article. They stated that they "did not provide any additional insights."
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The case-control study design was appropriate for the research question and applicable statistical methods were used (logistic regression models). However, the co-exposures were not accounted for in the analysis.
	Metric 13:	Statistical Power	Medium	The number of cases and controls was adequate to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis was sufficient to be reproducible with access to the analytical data.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described.

Additional Comments: For TCEP, dust levels in the home are used as surrogates for exposure, measured using MS/GC. PBDEs only were measured in serum (not organophosphate flame retardants). This case-control study used an adequate number of age- and gender-matched pairs. Multiple flame retardants were present in dust and in large ranges, depending on the flame retardant; adjustments were not made for multiple comparisons (separate models were assessed for each compound).

Overall Quality Determination High

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participant selection was well described, including timing of study, inclusion and exclusion criteria and case ascertainment.
	Metric 2:	Attrition	High	Supplemental Figure 1 provides detailed information and shows only 1 of 71 cases excluded.
	Metric 3:	Comparison Group	High	Cases and controls were recruited from the same Health Care Center during the same time period and were matched by age and gender.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Flame retardants were measured in house dust. Collection methods and analysis of dust was fully reported. TCEP concentrations in household dust were measured by GC/EI-MS.
	Metric 5:	Exposure Levels	Low	Only 2 levels of flame retardant exposure in the home were reported for cases and controls; TCEP concentrations above the median concentration were reported for both groups. Median TCEP concentrations were shown in the box plot in Fig. 1 but concentrations were not reported. Detection limits were not reported for organophosphate flame retardants.
	Metric 6:	Temporality	Medium	Exposure at the same residence for at least 2 years prior to diagnosis; however, it is unclear whether exposure duration was sufficient for thyroid cancer.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The thyroid cancer outcome was assessed by medical review of clinical and pathology data; pathological stage was assessed based on tumor size, location, metastasis.
	Metric 8:	Reporting Bias	Medium	Health outcomes and exposures are reported, except for median exposure levels of each of the flame retardants. Adjusted ORs with CIs reported for overall incidence as well as measures of tumor aggressiveness (i.e., pathologic stage). Some data were reported for PBDEs but not the organophosphate flame retardants, such as number of detects, detection limits.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariate adjustments were described (age, income, BMI). Regression analyses were adjusted for participant age and household income. BMI was both included and excluded as a covariate. Other potential confounders were considered, but not applied because they did not alter effect estimates (race, employment status, and smoking). Ionizing radiation exposure was considered but no participants reported prior exposure.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-thyroid cancer (papillary)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	A questionnaire was administered to participants, but the method for assessing covariates was not described, although the study indicated that protocols were approved by an institutional review board. Did not report if the questionnaire was validated or the types of data collected in the questionnaire, but there is no evidence of confounding.
	Metric 11:	Co-exposure Counfounding	Low	Many flame retardants were measured in household dust samples. The authors acknowledge and provide data on those most highly correlated. Statistical analyses modeled each flame retardant separately. The authors indicate that component analyses were conducted to systematically assess FR mixtures but did not provide them in the article. They stated that they "did not provide any additional insights."
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The case-control study design was appropriate for the research question and applicable statistical methods were used (logistic regression models). However, the co-exposures were not accounted for in the analysis.
	Metric 13:	Statistical Power	Medium	The number of cases and controls was adequate to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis was sufficient to be reproducible with access to the analytical data.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described.

Additional Comments: For TCEP, dust levels in the home are used as surrogates for exposure, measured using MS/GC. PBDEs only were measured in serum (not organophosphate flame retardants). This case-control study used an adequate number of age- and gender-matched pairs. Multiple flame retardants were present in dust and in large ranges, depending on the flame retardant; adjustments were not made for multiple comparisons (separate models were assessed for each compound).

Overall Quality Determination High

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participant selection was well described, including timing of study, inclusion and exclusion criteria and case ascertainment.
	Metric 2:	Attrition	High	Supplemental Figure 1 shows only 1 of 71 cases excluded and shows only 1 of 71 cases excluded.
	Metric 3:	Comparison Group	High	Cases and controls were recruited from the same Health Care Center during the same time period and were matched by age and gender.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Flame retardants were measured in house dust. Collection methods and analysis of dust was fully reported. TCEP concentrations in household dust were measured by GC/EI-MS.
	Metric 5:	Exposure Levels	Low	Levels of flame retardant exposure in the home were reported for cases and controls; TCEP concentrations above the median concentration were reported for both groups. Median TCEP concentrations were shown in the box plot in Fig. 1 but concentrations were not reported. Detection limits were not reported for organophosphate flame retardants.
	Metric 6:	Temporality	Medium	Exposure at the same residence for at least 2 years prior to diagnosis; however, it is unclear whether exposure duration was sufficient for thyroid cancer.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The thyroid cancer outcome was assessed by medical review of clinical and pathology data; pathological stage was assessed based on tumor size, location, metastasis.
	Metric 8:	Reporting Bias	High	Health outcomes and exposures are reported, except for median exposure levels of each of the flame retardants. Adjusted ORs with CIs reported for overall incidence as well as measures of tumor aggressiveness (i.e., pathologic stage). Some data were reported for PBDEs but not the organophosphate flame retardants, such as number of detects, detection limits.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariate adjustments were described (age, income, BMI). Regression analyses were adjusted for participant age and household income. BMI was both included and excluded as a covariate. Other potential confounders were considered, but not applied because they did not alter effect estimates (race, employment status, and smoking). Ionizing radiation exposure was considered but no participants reported prior exposure.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hoffman, K., Lorenzo, A., Butt, C. M., Hammel, S. C., Henderson, B. B., Roman, S. A., Scheri, R. P., Stapleton, H. M., Sosa, J. A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107:235-242.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	4161719

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	A questionnaire was administered to participants, but the method for assessing covariates was not described, although the study indicated that protocols were approved by an institutional review board. Did not report if the questionnaire was validated or the types of data collected in the questionnaire, but there is no evidence of confounding.
	Metric 11:	Co-exposure Counfounding	Low	Many flame retardants were measured in household dust samples. The authors acknowledge and provide data on those most highly correlated. Statistical analyses modeled each flame retardant separately. The authors indicate that component analyses were conducted to systematically assess FR mixtures but did not provide them in the article. They stated that they "did not provide any additional insights."
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The case-control study design was appropriate for the research question and applicable statistical methods were used (logistic regression models). However, the co-exposures were not accounted for in the analysis.
	Metric 13:	Statistical Power	Medium	The number of cases and controls was adequate to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis was sufficient to be reproducible with access to the analytical data.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described.

Additional Comments: For TCEP, dust levels in the home are used as surrogates for exposure, measured using MS/GC. PBDEs only ere measured in serum (not organophosphate flame retardants). This case-control study used an adequate number of age- and gender-matched pairs. Multiple flame retardants were present in dust and in large ranges, depending on the flame retardant; adjustments were not made for multiple comparisons (separate models were assessed for each compound).

Overall Quality Determination High

Study Citation:	Li, Y., Fu, Y., Hu, K., Zhang, Y., Chen, J., Zhang, S., Zhang, B., Liu, Y. (2020). Positive correlation between human exposure to organophosphate esters and gastrointestinal cancer in patients from Wuhan, China. Ecotoxicology and Environmental Safety 196:110548.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Gastrointestinal cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6747922

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Most key elements of study design and participation selection were described. Confirmed through pathology, cancer cases (n=74) were selected from a hospital in Wuhan, China. However, information on how controls (n=62) were selected was less transparent. It only mentions that they were "healthy enough to donate blood".
	Metric 2:	Attrition	Medium	There was minimal subject withdrawal from the study, and the outcome and exposure data seem to have been largely complete.
	Metric 3:	Comparison Group	Medium	There were some evidence that cases and controls were similar. For example, cases and controls on average had similar age and were recruited in the same time frame.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Quantitative measurement of chemicals/organophosphate esters were measured in plasma using HPLC-MS/MS and electrospray positive ionization methods.
	Metric 5:	Exposure Levels	Low	Range of exposure in the population were limited. Several chemicals, including TCEP, were not readily detected among cases and controls.
	Metric 6:	Temporality	Low	Temporality was established, exposures were assessed after cancer was diagnosed. However, it was unclear whether exposures, including TCEP, fall within relevant exposures windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Cases were assessed using appropriate methods; however, not enough information of how it was done is provided. Confirmed through pathology, cancer cases (n=74) were selected from a hospital in Wuhan, China.
	Metric 8:	Reporting Bias	Medium	An appropriate description of the outcome was reported in the abstract, introduction and methods. The odds ratio as the main effect estimates were reported with their respective confidence intervals (CI). Numbers of cases and controls were detailed for the logistic regressions, but they seemed to be very low based on the reported odds ratio and 95% CI.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was some adjustments for potential confounders in the logistic regression models. To test the main hypothesis, the authors adjusted for age, gender and stage, not for smoking and other potential confounders.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Li, Y., Fu, Y., Hu, K., Zhang, Y., Chen, J., Zhang, S., Zhang, B., Liu, Y. (2020). Positive correlation between human exposure to organophosphate esters and gastrointestinal cancer in patients from Wuhan, China. Ecotoxicology and Environmental Safety 196:110548.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Gastrointestinal cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6747922

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	A few potential confounders were obtained and assessed through clinical records. However, the authors do not mention why other important potential confounders were included in the study.
	Metric 11:	Co-exposure Counfounding	Low	There were direct evidence of potential co-exposures, but were not appropriately adjusted for in the statistical approaches and/or in the logistic regression models.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design chosen was appropriate for the research question to investigate the associations between TCEP exposures and gastrointestinal cancer. The study used a variety of statistical approaches, including logistic regression models, appropriate to address the research question.
	Metric 13:	Statistical Power	Medium	The number of participants (cases, n=74; controls, n=62) was adequate to detect an effect in the population; however, not adequate for a comprehensive subgroup analysis. Thus, this metric is borderline medium with a lean towards low.
	Metric 14:	Reproducibility of Analyses	Low	Description of statistical analyses is limited, thus it would be difficult to reproduce the authors' approach. Information about the treatment of missing values not provided, and the language discussing the Mann-Whitney U test is not entirely clear (unsure about "comparing the two groups" - cases and controls?).
	Metric 15:	Statistical Analysis	Low	The statistical models (logistic regression models) building process was appropriate and model assumptions were met (with respect to the outcome) to address the research question. However, the authors do not mentioned how they treated the independent variables and covariates in the models.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Biomarkers of exposure/organophosphate esters (including TCEP) were measured in plasma using HPLC-MS/MS and electrospray positive ionization methods.
	Metric 17:	Effect Biomarker	N/A	An effect biomarker was not assessed.
	Metric 18:	Method Sensitivity	Medium	Analytical methods to detect and measure biomarkers were sensitive.
	Metric 19:	Biomarker Stability	Medium	Some description of the sample collection and storage (at -20 degrees Celsius) until use for analytical purposes was provided.
	Metric 20:	Sample Contamination	High	Some information on sample handling to avoid or reduce potential contamination of the samples was described. Quality control measures were also taken into account to ensure reliable data.
	Metric 21:	Method Requirements	High	Information on instrumentation(s) that allowed for identification of the biomarkers, including TCEP, with a high degree of confidence and the required sensitivity was provided.
	Metric 22:	Matrix Adjustment	N/A	Matrix adjustment is not necessary for plasma biomarker samples.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Li, Y., Fu, Y., Hu, K., Zhang, Y., Chen, J., Zhang, S., Zhang, B., Liu, Y. (2020). Positive correlation between human exposure to organophosphate esters and gastrointestinal cancer in patients from Wuhan, China. Ecotoxicology and Environmental Safety 196:110548.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Gastrointestinal cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	6747922

Domain	Metric	Rating	Comments

Additional Comments: In this study, the authors examined the associations between investigated the associations between organophosphate esters, including TCEP, and the increase risk for gastrointestinal cancer in Wuhan, China. Overall, most of the methodology used in this study were apparently adequate. However, there were a number of chemicals that gave high frequencies of not detected values among the studied population and were not possible to be linked with the outcome of interest. In terms of the results, there was an increased risk for gastrointestinal cancer and exposure to organophosphate esters. Nonetheless, TCEP was not associated with gastrointestinal cancer due high frequencies of not detected values.

Overall Quality Determination Medium

Study Citation:	Liao, K., Zhao, Y., Qu, J., Yu, W., Hu, S., Fang, S., Zhao, M., Jin, H. (2023). Organophosphate esters concentrations in human serum and their associations with Sjögren syndrome. Environmental Pollution 331(Pt 1):121941.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Sjogren's syndrome
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364983

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	This study, conducted in China, analyzed associations between several organophosphate esters (OPEs) including tris (2-chloroethyl) phosphate [TCEP], and Sjogren's syndrome (SjS). The analysis sample included 145 healthy volunteers (controls) and 138 patients with SjS (cases) recruited from a hospital in Hangzhou, Zhejiang Province, China in 2021-2022. Information on participant selection was sparse. Recruitment methods and refusal rates were not described. There are concerns related to participant selection, but no direct evidence of selection bias.
	Metric 2:	Attrition	Medium	Of 318 participants recruited for the study, 35 participants (11%) were excluded due to missing demographic information or insufficient serum for OPE analysis. Attrition was not described by case status. However, the overall attrition of 11% is relatively low.
	Metric 3:	Comparison Group	Low	Controls were recruited from the same hospital as cases during the same time period. Controls were described as individuals "from the physical examination population," without much additional detail. Controls were matched to cases based on gender, which is appropriate and important because Sjogren's Syndrome is an autoimmune disease that is substantially more common in women than in men. A potential limitation is that matching wasn't also performed based on age. The mean age in cases was significantly higher than controls (61.46 $\pm$ 14.80 vs. 48.4$\pm$ 10.57 years); however, the statistical analyses adjusted for age. Eligibility criteria were not specified for controls. It is therefore uncertain whether controls were screened to exclude diagnoses of other conditions potentially associated with OPE exposure. This uncertainty, common to hospital-based case control studies, is a concern, but there is no direct evidence of bias.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liao, K., Zhao, Y., Qu, J., Yu, W., Hu, S., Fang, S., Zhao, M., Jin, H. (2023). Organophosphate esters concentrations in human serum and their associations with Sjögren syndrome. Environmental Pollution 331(Pt 1):121941.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Sjogren's syndrome
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364983

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	Exposure to several OPEs, including tris (2-chloroethyl) phosphate [TCEP], was assessed using serum samples. Samples were analyzed using high performance liquid chromatography with mass spectrometry, using internal standards and quality controls. The authors stated that no OPEs were detected in procedural blanks (ultra-pure water). The mean (SE) % recovery for TCEP was 100 ± 11. TCEP was detected in 95% of serum samples. Values below detection limits were replaced by the LOD divided by the square root of 2. There is limited data on the utility or validity of estimating TCEP exposure using serum samples. Issues such as the half-life of TCEP in serum, or the utility of estimating exposure using serum with or without including measures of TCEP metabolites, were not discussed.The authors did not adjust TCEP concentrations for serum lipid levels or adjust for lipids in statistical models. Serum TCEP concentrations in studies cited in the Discussion section were lipid-adjusted. However, OPE chemicals are often partly hydrophilic, so it is uncertain whether lipid adjustment is important for the analyses using serum measures of TCEP to estimate exposure.
	Metric 5:	Exposure Levels	Medium	TCEP exposure was analyzed using both a log transformed continuous variable and quartiles. TCEP was detectable in serum of 95% of participants. The mean was 0.54 ng/mL, and the range was from <LOD to 2.40 ng/mL in controls. The mean was 1.12 ng/mL and the range was <LOD to 8.69 ng/mL in cases. There was no evidence of insufficient variability to estimate associations.
	Metric 6:	Temporality	Low	There are substantial limitations and uncertainties in the temporality of exposure and outcome in this study. Serum TCEP was only assessed at a single timepoint, which didn't precede the development of the outcome. The timepoint at which TCEP exposure was measured might not have been appropriate to accurately reflect TCEP exposure and a potential latency period prior to the development of SjS. The paper did not provide information on time elapsed since diagnosis. Because exposure was measured after the outcome, reverse causation is a potential concern. Behavior changes, treatments, or disease-related changes in metabolism might have influenced concentrations of the serum TCEP exposure biomarker. The measurement of serum TCEP at only a single timepoint that didn't precede the determination of case status is an important limitation.

Domain 3: Outcome Assessment

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liao, K., Zhao, Y., Qu, J., Yu, W., Hu, S., Fang, S., Zhao, M., Jin, H. (2023). Organophosphate esters concentrations in human serum and their associations with Sjögren syndrome. Environmental Pollution 331(Pt 1):121941.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Sjogren's syndrome
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364983

Domain	Metric		Rating	Comments
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcome assessed in this study, Sjogren's syndrome (SjS), is a rare autoimmune disease. The authors described cases as diagnosed based on American-European Consensus Group classification criteria, which include the presence of auto-antibodies (SjS-related antigen A or B), or focal lymphocytic sialadenitis detected in salivary gland biopsies. They stated that these criteria have high sensitivity (93.5%) and specificity (94%). The study did not discuss whether the cases included were limited to primary SjS, or whether a proportion had secondary SjS (i.e. associated with another underlying autoimmune disorder such as rheumatoid arthritis). There was no mention of assessing autoantibodies in controls or screening controls for either SjS or other autoimmune disorders.
	Metric 8:	Reporting Bias	High	Results for associations described in the Materials and Methods section of the paper were presented. Effect estimates were reported with measures of variability (confidence intervals).
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Controls were matched to cases based on gender during participant selection because Sjogren's occurs disproportionately in women, but the statistical methods didn't include matched pairs analyses. Conditional logistic regression could have been used to account for the matching, but the study appears to have used ordinary logistic regression and gender was omitted from the pooled model. However, some analyses were stratified by gender. The authors considered other potential confounders using exclusion by single-factor screening to develop a final model that included age, smoking, and alcohol consumption. Other potential confounders (BMI, education, household income, parity, and diet) had similar distributions in cases and controls and were excluded from the final models. Residual confounding cannot be ruled out, but there was no evidence of substantial confounding bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were assessed using questionnaires. The validity of the questionnaires used to measure confounders was not discussed. However, there was no evidence that covariate measures had poor validity.
	Metric 11:	Co-exposure Counfounding	Medium	TCEP was not strongly correlated with other OPE chemicals measured in this study (Spearman's R <0.50). Joint effects of multiple OPE chemicals were also examined in a supplementary mixtures analysis using weighted quantile sum regression. Potential confounding by non-OPE co-exposures is possible. The authors mentioned infectious agents, UV radiation, and drugs as potential SjS triggers, but didn't assess these exposures in participants.
Domain 5: Analysis				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liao, K., Zhao, Y., Qu, J., Yu, W., Hu, S., Fang, S., Zhao, M., Jin, H. (2023). Organophosphate esters concentrations in human serum and their associations with Sjögren syndrome. Environmental Pollution 331(Pt 1):121941.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Sjogren's syndrome
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364983

Domain	Metric		Rating	Comments
	Metric 12:	Study Design and Methods	High	The authors only made a few mentions of the study type, and were inconsistent in the study type specified. In the abstract and in one sentence in the paper, the authors state that this is a cross-sectional study. However, in another sentence in the paper, the authors describe the "case-control population" and in another sentence the authors refer to the limitations of their "case-cohort design". Despite these issues with terminology and some lacking details, the methods described in the paper are generally consistent with a case-control design, which would be appropriate because Sjogren's syndrome is a rare disease. Logistic regression was used, which is appropriate for a case-control study.
	Metric 13:	Statistical Power	Medium	This study sample included 145 controls and 138 cases, and there was variability in serum TCEP concentrations (median, IQR = 0.45, 0.26-0.76 ng/mL). However, the power to detect associations was lower in sex-stratified analyses (males: 45 cases, 51 controls; females 93 cases, 94 controls). The authors mentioned the relatively small study sample size as a potential limitation.
	Metric 14:	Reproducibility of Analyses	Medium	The statistical analysis methods were clearly described and appear to be readily reproducible.
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for case and control characteristics. Logistic regression models were used to analyze the odds of Sjogren's syndrome associated with exposure to TCEP and other OPE chemicals. Crude and adjusted odds ratios with 95% confidence intervals and p-values were shown. In the overall sample, TCEP exposure was analyzed both using quartiles and using a log transformed continuous variable. Results using quartiles suggested that the association was non-linear and U-shaped; the utility of sex-stratified analyses that used only a continuous exposure variable may be limited.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	TCEP concentrations in this study were measured in participant serum. The authors did not cite evidence to support the validity of serum TCEP as a measure of exposure. However, data on the use of serum vs other media (e.g. whole blood, urine, hair) to estimate OPE exposures is limited at this time. This study did not measure metabolites of TCEP [e.g., bis(2-chloroethyl) phosphate (BCEP)] along with the parent chemical. It is uncertain whether including metabolites along with the parent compound would meaningfully influence TCEP exposure estimation based on serum measures.
	Metric 17:	Effect Biomarker	N/A	Although biomarkers of effect (autoantibodies) are one of the criteria for SjS diagnosis, the biomarker metrics weren't evaluated for biomarkers of effect for this study because additional non-biomarker components are included in the diagnostic criteria, and autoantibody biomarkers weren't assessed in controls.
	Metric 18:	Method Sensitivity	Medium	TCEP was detected in 95% of serum samples from controls and 84% of samples from cases.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liao, K., Zhao, Y., Qu, J., Yu, W., Hu, S., Fang, S., Zhao, M., Jin, H. (2023). Organophosphate esters concentrations in human serum and their associations with Sjögren syndrome. Environmental Pollution 331(Pt 1):121941.
Health Outcome(s) and Reported Health Effect(s):	Immune/Hematological-Sjogren's syndrome
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364983

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Substantial details about sample storage weren't discussed. However, there is evidence that the timeframe of sample storage was relatively short based on participants in the study being recruited in 2021-2022 and the study being published in 2023.
	Metric 20:	Sample Contamination	Medium	The authors stated that procedural blanks had no detectable levels of TCEP. However, there was limited information on sample collection and handling protocols.
	Metric 21:	Method Requirements	Medium	The study reported using HPLC-MS with internal standards and quality controls. The mean (SE) % recovery reported for matrix spiked TCEP samples was 100 $\pm$ 11. There was no evidence to suggest important deficiencies in measurement methods.
	Metric 22:	Matrix Adjustment	Medium	The authors did not adjust TCEP concentrations for serum lipid levels or adjust for lipids in statistical models. In the discussion section, serum TCEP concentrations described in other studies were lipid-adjusted. However, as OPE chemicals are often partly hydrophilic, it is unclear whether lipid adjustment is important for analyses using serum measures of TCEP.

Additional Comments: This study assessed the association between serum TCEP concentrations and Sjögren's syndrome (SjS) in a study of 138 SjS cases and 145 controls in Hangzhou, China. Increased serum levels of a combination of several organophosphate esters were found to be associated with an increased odds of SjS. However, the only statistically significant associations found between TCEP serum levels and SjS were for an analysis using exposure quartiles in which the observed relationship with SjS was non-linear. Compared to the lowest TCEP exposure quartile (first quartile), the odds of SjS were significantly lower in the second and third quartile and higher in the fourth quartile in the crude analyses. After adjusting for potential confounders, the inverse associations observed for the second and third quartile remained statistically significant, but the positive association for the fourth quartile was attenuated and no longer statistically significant (p = 0.143). When TCEP was log-transformed and treated as a continuous variable, no statistically significant increase in odds of SjS was found in the crude or adjusted models. Strengths of this study include a clinical assessment of SjS that was reported to have high specificity and sensitivity. However, the authors did not specify whether cases were limited to primary SjS vs both primary and secondary disease. Other limitations include a lack of details on potential temporal changes in exposure, lack of detailed health information on the controls and lack of description of diagnosis date and treatments for the cases. The potential for residual confounding of the assessed associations cannot be ruled out. Sjogren's syndrome occurs disproportionately in females. Although controls were matched to cases based on gender during participant selection, the statistical analyses didn't account for this matching, and some but not all analyses were stratified based on gender. Controls and cases weren't matched by age, but the analyses adjusted for age, and other potential confounders were also considered in modeling. Some aspects of the study design terminology and methods could have been discussed with more clarity and detail. There are substantial temporality concerns relevant to interpreting the findings of this study, particularly the limitation that the exposure (serum TCEP levels) was only assessed at a single timepoint, which didn't precede the development of the outcome (SjS status).

Overall Quality Determination Medium

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-Papillary thyroid cancer; Thyroid hormone levels: triiodothyronine (T3) free T3, thyroxine (T4), free FT4, thyroid stimulating hormone (TSH)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Cross-sectional study examining exposure to flame retardant chemicals and associations with papillary thyroid cancer, and thyroid hormones among controls. Initially, 350 patients with papillary thyroid cancer were recruited, along with 784 controls, from a hospital in Shandong Province between 2020 and 2021. The analysis sample included 242 cases and 239 controls. Eligibility criteria were age >18 years, more than 6 months residing in the area, sufficient serum for analysis, and availability of triglyceride and cholesterol measures. Information on several aspects of participant selection was limited, including the number of individuals excluded for each reason. Recruitment methods and refusal rates were not described. The study did not describe how cases were identified and selected, did not discuss time elapsed since initial cancer diagnosis, and did not mention whether cases had other diagnoses (i.e., health status might have influenced TCEP exposure-related behaviors such as dietary habits). Indeed, 12 cases had diabetes. There are several concerns related to participant selection, but no direct evidence of selection bias.
	Metric 2:	Attrition	Low	The study initially recruited 350 thyroid cancer patients and 784 controls. After exclusions based on eligibility, the sample included 242 (69.1%) cases and 239 (30.5%) controls. It was unclear why participants, in particular a high proportion of controls, were excluded: attrition due to specific eligibility criteria was not described. Though the lack of information is a concern, there was no evidence of bias associated with exclusions of either cases or controls.
	Metric 3:	Comparison Group	Low	Controls were recruited from the same hospital as cases, during the same time period. Controls were randomly selected individuals receiving a physical exam, matched to cases on age (within 5 years) and sex, and screened to ensure healthy thyroid function (including ultrasounds). Eligibility criteria for controls was the same as cases. Information on several aspects of control selection was limited. The study did not describe whether physical exams used to identify controls were routine well checks, and there was limited information on health status. Because two controls and 12 cases had diabetes, models adjusted for diabetes, but diagnoses with other conditions is uncertain. There are several concerns related to control selection, but no direct evidence of bias.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-Papillary thyroid cancer; Thyroid hormone levels: triiodothyronine (T3) free T3, thyroxine (T4), free FT4, thyroid stimulating hormone (TSH)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	A single serum sample was used to measure levels of nine OPEs, including tris (2-chloroethyl) phosphate [TCEP], in all study participants. Samples were analyzed using gas chromatography−triple quadrupole mass spectrometry with standards and quality controls that included procedural blanks. The mean (SE) % recovery for TCEP during calibration was 119 ± 12.4; TCEP was detected in 64% of cases and 74% of controls. Values below method detection limits (MDL) were replaced by MDL/2. Use of serum samples to estimate TCEP exposure is not well-established, and issues such as half-life were not discussed. However, there is no evidence that measures were not appropriate.
	Metric 5:	Exposure Levels	Medium	TCEP exposure was analyzed using approximate exposure quartiles, or natural log transformed continuous variables. Levels were detectable in 74% of controls and 64% of cases. Though concentrations were low, there was no evidence of insufficient variability to estimate associations.
	Metric 6:	Temporality	Low	As is common with case control studies of cancer outcomes, temporality of exposure and outcome is uncertain. There was no information on the timing of sample collection relative to date of diagnosis: disease-related changes in metabolism, cancer treatments, or shifts in exposure-related behaviors among cases might have affected TCEP measures. In addition, residential mobility among participants beyond the 6-month eligibility requirement was not described. Changes in residence might affect whether current measures reflect exposure levels in indoor environments during the period relevant for cancer genesis. There were several concerns, but no direct evidence that exposure levels were influenced by factors such as changes in metabolism, by cancer treatments, or shifts in behavior.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes included diagnosed papillary thyroid cancer and serum levels of several thyroid hormones. (1) Thyroid cancer: Cases were recruited from a hospital and described as having diagnosed cancer. An important strength is that controls were screened to ensure normal thyroid function, though criteria used were not specified. Including a single type of thyroid cancer was also a strength. However, heterogeneity in disease stage or tumor characteristics, potentially related to etiology, was not discussed. As noted earlier, there was no discussion of time since cancer diagnosis. (2) Thyroid hormones: Levels of thyroid hormones were measured in serum samples collected from controls using an electro-chemiluminescence immunoassay analyzer. Cancer cases were not included in these analyses. Measures included triiodothyronine (T3) free T3, thyroxine (T4), free FT4, and thyroid stimulating hormone (TSH). The distribution of hormone levels indicated substantial variability, and that a few participants had values slightly outside of reference ranges. Nonetheless, there was no evidence that this issue would meaningfully influence results.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-Papillary thyroid cancer; Thyroid hormone levels: triiodothyronine (T3) free T3, thyroxine (T4), free FT4, thyroid stimulating hormone (TSH)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed as described in the study aims. Results are reported with measures of variability. However, the number of cases included was not characterized.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Models analyzing TCEP and thyroid cancer adjusted for age, sex, BMI, current smoking status, current alcohol drinking, and diabetes status. Models analyzing thyroid hormones among controls adjusted for the same variables. Confounders appear to have been selected a priori. Residual confounding by factors such as past smoking or alcohol use, weight or dietary changes, and socioeconomic factors is a potential concern, but there is no evidence of bias.
	Metric 10:	Covariate Characterization	Medium	The validity of questionnaires used to measure confounders were not discussed. However, there was no evidence that covariate measures had poor validity.
	Metric 11:	Co-exposure Counfounding	Medium	Numerous OPEs were measured in participant serum samples. Weighted quantile sum regression was used to analyze associations with a chemical mixture including co-exposures. Moreover, TCEP was not strongly correlated with other chemicals measured in this study.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This case-control study in China analyzed associations between thyroid cancer and several organophosphate esters (OPEs) and flame retardants (FRs), including tris (2-chloroethyl) phosphate [TCEP]. The use of a case control study design was appropriate for the research question given that thyroid cancer is a rare disease. A crosss-sectional analysis was also appropriate for evaluating how TCEP exposure was associated with thyroid hormone levels among adults with normal thyroid function.
	Metric 13:	Statistical Power	Medium	The analysis sample included 242 (87 male, 155 female) cases and 239 (88 male, 151 female) controls. Concentrations of TCEP were low, but there was variability in exposure. Analysis of thyroid hormone levels were limited to controls, but power was increased by the use of continuous outcome measures. Power to detect associations was lower in sex-stratified than in pooled analyses. However, there was no evidence of inadequate sensitivity.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and readily reproducible.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-Papillary thyroid cancer; Thyroid hormone levels: triiodothyronine (T3) free T3, thyroxine (T4), free FT4, thyroid stimulating hormone (TSH)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for case and control characteristics, and the distribution of thyroid hormones was shown. Logistic regression models that included matching variables were used to analyze thyroid cancer, and linear regression was used to analyze thyroid hormone outcomes. TCEP was analyzed using quartiles or natural log transformed continuous variables. Thyroid hormone outcomes were also log transformed in linear regression models.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	TCEP concentrations were measured in patient serum. As there is limited data on the use of serum to estimate of OPE exposures at this time, the authors did not cite evidence to directly support the validity of serum TCEP as a measure of exposure. However, they cited other studies in China with comparable levels of TCEP in serum, as well as studies with considerably higher levels. Given the short half-life of OPE metabolites in urine, the authors suggested use of serum may be an advantage as this medium may reflect a longer exposure time window but did not cite supporting evidence of utility of a single serum sample to quantify exposure. A potential limitation is that metabolites of TCEP [major metabolite bis(2-chloroethyl) phosphate (BCEP)] were not assayed in addition to the parent chemical. There is however no evidence that exposure rankings based on metabolites vs the parent compound might differ. Despite potential concerns, there was no evidence that measures had poor validity, and no evidence of either substantial exposure misclassification or bias.
	Metric 17:	Effect Biomarker	High	Serum measures of thyroid hormones were analyzed to estimate associations between TCEP and thyroid function in healthy adults (controls). A strength of the approach was measurement of both total and free levels of T3 and T4, along with thyroid stimulating hormone (TSH). Thyroid hormones are typically used as diagnostic criteria for thyroid conditions. However, outcomes were based on a single serum measure in a sample of moderate size. Assay details were not provided, but a standard clinical method and analyzer was used.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure and was low: $8.0 \times 10{-}2$ ng/mL. Detection limits for thyroid hormones were not described, but there was no evidence of deficiencies.
	Metric 19:	Biomarker Stability	Medium	Samples were not likely to have been stored for a lengthy period. There is no data on stability, but no evidence of degradation.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Endocrine-Papillary thyroid cancer; Thyroid hormone levels: triiodothyronine (T3) free T3, thyroxine (T4), free FT4, thyroid stimulating hormone (TSH)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 20:	Sample Contamination	Medium	Lab procedures including precleaning glass containers using pesticide-grade solvents. The study does not discuss whether contamination was an issue of concern for TCEP or discuss detection of TCEP in procedural blanks. However, MDL for TCEP was low ($8.0 \times 10-2$ ng/mL) and there was a moderate proportion of samples with levels below MDL (26% in controls, 36% in cases). There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	Medium	The laboratory methods described appeared to be appropriate. Samples were analyzed using gas chromatography−triple quadrupole mass spectrometry with reference standards for calibration, and the use of quality controls that included procedural blanks. The mean (SE) % recovery for TCEP during calibration was 119 ± 12.4.
	Metric 22:	Matrix Adjustment	Medium	TCEP exposure variables were analyzed adjusted for lipid weight, as in numerous other studies. Total cholesterol and triglycerides were used for this adjustment.

Additional Comments: This hospital-based case control (242 cases, 239 controls) study in Shandong, China analyzed the relationship between serum levels of tris (2-chloroethyl) phosphate [TCEP] and papillary thyroid cancer. The study found an association between TCEP and thyroid cancer that was negative overall, with null and non-linear associations after stratifying by sex. The authors cited evidence of potential sex-specific effects on thyroid function of the chemicals examined in this study. Strengths of this analysis include screening to ensure that controls were cancer free and had healthy thyroid function. Important concerns include limited detail on the health status of controls, presence of other diagnoses in cases (12 had diabetes), and the use of prevalent thyroid cancer cases without discussion of diagnosis dates, treatments, or behavior changes that might have affected organophosphate chemical metabolism. Exposure measurements were measured cross-sectionally, which limits interpretations of temporality. The study also analyzed the cross-sectional association between serum TCEP and levels of thyroid hormones among controls. Controls were free of thyroid disease, but eligibility criteria did not clearly specify excluding other diagnoses. Two controls had diabetes, which was included as a covariate in the analyses. Consistent with hypothyroid activity, TCEP was negatively associated with T3, free T3, T4, and free T4 and positively associated with TCEP in females ($p<0.05$ except for total T4). Among male participants and in the overall sample, there was also a significant ($p<0.05$) negative association between TCEP and free T4. In contrast to females, however, there was a non-significant negative association with TSH among men. The authors stated that the use of serum to estimate TCEP exposure in this study may be an advantage relative to studies that measure short-lived TCEP metabolites in urine samples. However, there is limited evidence on the validity of biomarkers of TCEP exposure, and TCEP metabolites were not measured in serum along with the parent compound.

Overall Quality Determination Medium

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Cross-sectional study examining exposure to flame retardant chemicals and associations with papillary thyroid cancer, and thyroid hormones among controls. Initially, 350 patients with papillary thyroid cancer were recruited, along with 784 controls, from a hospital in Shandong Province between 2020 and 2021. The analysis sample included 242 cases and 239 controls. Eligibility criteria were age >18 years, more than 6 months residing in the area, sufficient serum for analysis, and availability of triglyceride and cholesterol measures. Information on several aspects of participant selection was limited, including the number of individuals excluded for each reason. Recruitment methods and refusal rates were not described. The study did not describe how cases were identified and selected, did not discuss time elapsed since initial cancer diagnosis, and did not mention whether cases had other diagnoses (i.e., health status might have influenced TCEP exposure-related behaviors such as dietary habits). Indeed, 12 cases had diabetes. There are several concerns related to participant selection, but no direct evidence of selection bias.
	Metric 2:	Attrition	Low	The study initially recruited 350 thyroid cancer patients and 784 controls. After exclusions based on eligibility, the sample included 242 (69.1%) cases and 239 (30.5%) controls. It was unclear why participants, in particular a high proportion of controls, were excluded: attrition due to specific eligibility criteria was not described. Though the lack of information is a concern, there was no evidence of bias associated with exclusions of either cases or controls.
	Metric 3:	Comparison Group	Low	Controls were recruited from the same hospital as cases, during the same time period. Controls were randomly selected individuals receiving a physical exam, matched to cases on age (within 5 years) and sex, and screened to ensure healthy thyroid function (including ultrasounds). Eligibility criteria for controls was the same as cases. Information on several aspects of control selection was limited. The study did not describe whether physical exams used to identify controls were routine well checks, and there was limited information on health status. Because two controls and 12 cases had diabetes, models adjusted for diabetes, but diagnoses with other conditions is uncertain. There are several concerns related to control selection, but no direct evidence of bias.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	A single serum sample was used to measure levels of nine OPEs, including tris (2-chloroethyl) phosphate [TCEP], in all study participants. Samples were analyzed using gas chromatography−triple quadrupole mass spectrometry with standards and quality controls that included procedural blanks. The mean (SE) % recovery for TCEP during calibration was 119 ± 12.4; TCEP was detected in 64% of cases and 74% of controls. Values below method detection limits (MDL) were replaced by MDL/2. Use of serum samples to estimate TCEP exposure is not well-established, and issues such as half-life were not discussed. However, there is no evidence that measures were not appropriate.
	Metric 5:	Exposure Levels	Medium	TCEP exposure was analyzed using approximate exposure quartiles, or natural log transformed continuous variables. Levels were detectable in 74% of controls and 64% of cases. Though concentrations were low, there was no evidence of insufficient variability to estimate associations.
	Metric 6:	Temporality	Low	As is common with case control studies of cancer outcomes, temporality of exposure and outcome is uncertain. There was no information on the timing of sample collection relative to date of diagnosis: disease-related changes in metabolism, cancer treatments, or shifts in exposure-related behaviors among cases might have affected TCEP measures. In addition, residential mobility among participants beyond the 6-month eligibility requirement was not described. Changes in residence might affect whether current measures reflect exposure levels in indoor environments during the period relevant for cancer genesis. There were several concerns, but no direct evidence that exposure levels were influenced by factors such as changes in metabolism, by cancer treatments, or shifts in behavior.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes included diagnosed papillary thyroid cancer. Cases were recruited from a hospital and described as having diagnosed cancer. An important strength is that controls were screened to ensure normal thyroid function, though criteria used were not specified. Including a single type of thyroid cancer was also a strength. However, heterogeneity in disease stage or tumor characteristics, potentially related to etiology, was not discussed. As noted earlier, there was no discussion of time since cancer diagnosis.
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed as described in the study aims. Results are reported with measures of variability. However, the number of cases included was not characterized.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Medium	Models analyzing TCEP and thyroid cancer adjusted for age, sex, BMI, current smoking status, current alcohol drinking, and diabetes status. Models analyzing thyroid hormones among controls adjusted for the same variables. Confounders appear to have been selected a priori. Residual confounding by factors such as past smoking or alcohol use, weight or dietary changes, and socioeconomic factors is a potential concern, but there is no evidence of bias.
	Metric 10:	Covariate Characterization	Medium	The validity of questionnaires used to measure confounders were not discussed. However, there was no evidence that covariate measures had poor validity.
	Metric 11:	Co-exposure Counfounding	Medium	Numerous OPEs were measured in participant serum samples. Weighted quantile sum regression was used to analyze associations with a chemical mixture including co-exposures. Moreover, TCEP was not strongly correlated with other chemicals measured in this study.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This case-control study in China analyzed associations between thyroid cancer and several organophosphate esters (OPEs) and flame retardants (FRs), including tris (2-chloroethyl) phosphate [TCEP]. The use of a case control study design was appropriate for the research question given that thyroid cancer is a rare disease. A crosss-sectional analysis was also appropriate for evaluating how TCEP exposure was associated with thyroid hormone levels among adults with normal thyroid function.
	Metric 13:	Statistical Power	Medium	The analysis sample included 242 (87 male, 155 female) cases and 239 (88 male, 151 female) controls. Concentrations of TCEP were low, but there was variability in exposure. Analysis of thyroid hormone levels were limited to controls, but power was increased by the use of continuous outcome measures. Power to detect associations was lower in sex-stratified than in pooled analyses. However, there was no evidence of inadequate sensitivity.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and readily reproducible.
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for case and control characteristics, and the distribution of thyroid hormones was shown. Logistic regression models that included matching variables were used to analyze thyroid cancer, and linear regression was used to analyze thyroid hormone outcomes. TCEP was analyzed using quartiles or natural log transformed continuous variables. Thyroid hormone outcomes were also log transformed in linear regression models.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric		Rating	Comments
	Metric 16:	Use of Biomarker of Exposure	Medium	TCEP concentrations were measured in patient serum. As there is limited data on the use of serum to estimate of OPE exposures at this time, the authors did not cite evidence to directly support the validity of serum TCEP as a measure of exposure. However, they cited other studies in China with comparable levels of TCEP in serum, as well as studies with considerably higher levels. Given the short half-life of OPE metabolites in urine, the authors suggested use of serum may be an advantage as this medium may reflect a longer exposure time window but did not cite supporting evidence of utility of a single serum sample to quantify exposure. A potential limitation is that metabolites of TCEP [major metabolite bis(2-chloroethyl) phosphate (BCEP)] were not assayed in addition to the parent chemical. There is however no evidence that exposure rankings based on metabolites vs the parent compound might differ. Despite potential concerns, there was no evidence that measures had poor validity, and no evidence of either substantial exposure misclassification or bias.
	Metric 17:	Effect Biomarker	High	Serum measures of thyroid hormones were analyzed to estimate associations between TCEP and thyroid function in healthy adults (controls). A strength of the approach was measurement of both total and free levels of T3 and T4, along with thyroid stimulating hormone (TSH). Thyroid hormones are typically used as diagnostic criteria for thyroid conditions. However, outcomes were based on a single serum measure in a sample of moderate size. Assay details were not provided, but a standard clinical method and analyzer was used.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure and was low: $8.0 \times 10{-}2$ ng/mL. Detection limits for thyroid hormones were not described, but there was no evidence of deficiencies.
	Metric 19:	Biomarker Stability	Medium	Samples were not likely to have been stored for a lengthy period. There is no data on stability, but no evidence of degradation.
	Metric 20:	Sample Contamination	Medium	Lab procedures including precleaning glass containers using pesticide-grade solvents. The study does not discuss whether contamination was an issue of concern for TCEP or discuss detection of TCEP in procedural blanks. However, MDL for TCEP was low ($8.0 \times 10{-}2$ ng/mL) and there was a moderate proportion of samples with levels below MDL (26% in controls, 36% in cases). There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	Medium	The laboratory methods described appeared to be appropriate. Samples were analyzed using gas chromatography−triple quadrupole mass spectrometry with reference standards for calibration, and the use of quality controls that included procedural blanks. The mean (SE) % recovery for TCEP during calibration was 119 ± 12.4.
	Metric 22:	Matrix Adjustment	Medium	TCEP exposure variables were analyzed adjusted for lipid weight, as in numerous other studies. Total cholesterol and triglycerides were used for this adjustment.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, M., Li, A., Meng, L., Zhang, G., Guan, X., Zhu, J., Li, Y., Zhang, Q., Jiang, G. (2022). Exposure to Novel Brominated Flame Retardants and Organophosphate Esters and Associations with Thyroid Cancer Risk: A Case-Control Study in Eastern China. Environmental Science & Technology 56(24):17825-17835.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Papillary thyroid cancer
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364830

Domain	Metric	Rating	Comments
Additional Comments:	This hospital-based case control (242 cases, 239 controls) study in Shandong, China analyzed the relationship between serum levels of tris (2-chloroethyl) phosphate [TCEP] and papillary thyroid cancer. The study found an association between TCEP and thyroid cancer that was inverse overall, with null and non-linear associations after stratifying by sex. The authors cited evidence of potential sex-specific effects on thyroid function of the chemicals examined in this study. Strengths of this analysis include screening to ensure that controls were cancer free and had healthy thyroid function. Important concerns include limited detail on the health status of controls, presence of other diagnoses in cases (12 had diabetes), and the use of prevalent thyroid cancer cases without discussion of diagnosis dates, treatments, or behavior changes that might have affected organophosphate chemical metabolism. Exposure measurements were measured cross-sectionally, which limits interpretations of temporality. The authors stated that the use of serum to estimate TCEP exposure in this study may be an advantage relative to studies that measure short-lived TCEP metabolites in urine samples. However, there is limited evidence on the validity of biomarkers of TCEP exposure, and TCEP metabolites were not measured in serum along with the parent compound.		

Overall Quality Determination Medium

Study Citation:	Liu, Y., Li, Y., Dong, S., Han, L., Guo, R., Fu, Y., Zhang, S., Chen, J. (2021). The risk and impact of organophosphate esters on the development of female-specific cancers: Comparative analysis of patients with benign and malignant tumors. Journal of Hazardous Materials 404(Pt B):124020.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Benign breast tumor; breast cancer; benign tumor of the uterus; cervical cancer.
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	7537904

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Study population recruited from a 3A hospital in Wuhan, China. Number of participants reported but participation rate not provided. The study included female patients with tumor/cancer disease but other inclusion and exclusion criteria not specified. Cancer cases was confirmed by two independent pathologists after histopathological analysis.
	Metric 2:	Attrition	High	No subject was reported withdrawal from the study. The outcome data and exposure measurement were complete for the study participants.
	Metric 3:	Comparison Group	Low	The study recruited patients with malignant tumor (cancer cases) and benign tumor. Comparison was made between cancer and benign tumor patients only. The control group of non-tumor population was not recruited. The similarity between benign and cancer groups is a concern, and lack of description of similarity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure chemicals were extracted from plasma samples using liquid-liquid extraction method, and directly measured using HPLC-MS/MS.
	Metric 5:	Exposure Levels	Medium	Continuous exposure chemical concentrations were measured for each patient. The range is adequate to detect the exposure-outcome association.
	Metric 6:	Temporality	Low	Due to the nature of cross-sectional design of this study, temporality is not well-established. The study only reported age of participants, other temporality supporting information was not provided. It is unclear whether exposures fall within relevant exposure windows for the outcomes.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Cancer diagnosis performed using postoperative histopathological analysis of tumors, then confirmed by 2 independent pathologists. The measurement of outcomes has high level of certainty.
	Metric 8:	Reporting Bias	High	Measured outcomes were reported in the methods section. Number of subjects in each case group and reference group provided. The effect estimates were reported as ORs with 95% CI in both adjusted and unadjusted model. Spearman correlations of each chemical in each group were reported. The results are fully tabulated in this study.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Liu, Y., Li, Y., Dong, S., Han, L., Guo, R., Fu, Y., Zhang, S., Chen, J. (2021). The risk and impact of organophosphate esters on the development of female-specific cancers: Comparative analysis of patients with benign and malignant tumors. Journal of Hazardous Materials 404(Pt B):124020.
Health Outcome(s) and Reported Health Effect(s):	Cancer/Carcinogenesis-Benign breast tumor; breast cancer; benign tumor of the uterus; cervical cancer.
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	7537904

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Medium	The covariates include age, hypertension, liver function, urine routine tests results, and renal function. Covariates were adjusted in the adjusted model. Distribution of covariates among study population was provided, but a description of how covariates were selected into the model was not reported.
	Metric 10:	Covariate Characterization	Medium	The covariate assessment methods or validation is not provided, but there is little to no evidence to indicate poor validity.
	Metric 11:	Co-exposure Counfounding	Low	Co-exposure of other OPE chemicals was measured. But the statistical model didn't make appropriate adjustment for co-exposure chemicals when evaluating the association between target chemical and outcomes.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	The cross-sectional design of this study to evaluate the association between OPE concentrations and cancer status is not appropriate, since cancer is usually a long-term disease. The statistical methods used to assess the correlation of chemicals and exposure-outcome association is appropriate.
	Metric 13:	Statistical Power	Low	Overall, 258 participants included in this study but divided into 4 subgroups based on tumor/cancer type. The number of participants in each subgroups ranged from 45 to 78. The number of participants are limited so it's a concern if there is sufficient statistical power to detect the effect of interest.
	Metric 14:	Reproducibility of Analyses	Medium	Sufficient information of study design and analysis were provided. The description is adequate to understand what has been done and conceptually reproducible with the access of original data.
	Metric 15:	Statistical Analysis	High	The binary logistic regression model to calculate odd ratios is appropriate. Lack of description of inclusion and exclusion of variables in adjusted model is a concern, but overall the analysis is transparent and appropriate.

Additional Comments: Overall, this cross-sectional study evaluated the association between organophosphate esters (OPEs) concentrations and female-specific tumor, and compared the OPE levels in benign and malignant disease patients. The exposure and outcome were measured using well-established methods with high certainty, and the analysis methods is overall appropriate. However, some concerns exist and could downgrade this study. The number of participants is small indicates potential lack of statistical power to detect the effect of interest. The cross-sectional design is not appropriate since cancer is usually not a acute disease, reverse causality could be an issue for this study. No health or non-cancer participants as reference group or control, so similarity of groups is a concern.

Overall Quality Determination Low

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants in this prospective birth cohort study were pulled from the Health Outcomes and Measures of the Environment (HOME) study. Pregnant women were recruited from March 2003-January 2006 in the greater Cincinnati, Ohio metropolitan area. Details of the HOME study are available in Braun et al. 2017 (HEROID: 6749104). In Braun et al. 2017 inclusion criteria and recruitment are extensively detailed and there is no evidence to suggest selection bias occurred in the original creation of the HOME study401 mother-child pairs were included in the original HOME study. Inclusion criteria for the analysis in this study were: having data on dust and urinary organophosphate esters and replacement brominated flame retardants, having data on covariates, and data on respiratory outcomes through follow-up until 5 years. 342 mother-child pairs were included in the present study.In general, there is no significant evidence for selection bias. However, no data is presented comparing those included in the study to participants who were excluded after originally participating in the HOME study.
	Metric 2:	Attrition	Medium	There were n=59 pregnant mother-child pair participants out of 401 who were included in the original HOME study but not the present analysis; all exclusion reasons were related to missing data, but exact breakdowns are not provided (ex; how many were excluded due to missing covariate data vs. missing exposure or outcome data). However, there is no reason to suggest that exclusion of those participants was inappropriate.
	Metric 3:	Comparison Group	High	Participants were all recruited from the same eligible population of pregnant women in the Greater Cincinnati, Ohio metropolitan area. The study presents a table of demographic characteristics split by exposure to all relevant exposures, including TCEP. Significant differences were reported by race/ethnicity and serum cotinine (significant differences were reported additionally for serum cotinine when assessing BCEP). All these factors were adjusted for in statistical analysis.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Exposure to TCEP was measured in dust samples. Dust was sampled from a 1-m squared area floor area for median sampling time of 275 seconds in the ”main activity room” of participants home at around 20 weeks gestation. TCEP was quantified using high-performance liquid chromatography mass spectrometry.
	Metric 5:	Exposure Levels	Medium	Exposure ranges were likely large enough to allow an exposure-response estimate. For example, TCEP in dust (5th-95th percentile): 0.06 - 9.31 ug/g dust.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Medium	Exposure was measured in the homes of mothers during pregnancy, and outcomes were measured every 6 months after birth. Temporality is thus established, but it is uncertain whether or not dust concentrations of TCEP during development represent a etiologically relevant time period for respiratory outcomes of children.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Respiratory outcomes included wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), and peak expiratory flow (PEF). Wheeze, respiratory infections, and hay fever or allergies were measured busting standardized questions, based on NHANES questionnaires. Questionnaires were provided every 6 months until the age of 5 years old. FEV1 and PEF were measured at age 5 years old using a portable spirometer, and measures were reported to follow the American Thoracic Society criteria and conducted by trained research assistants. In general, there is no significant evidence of outcome misclassification.
	Metric 8:	Reporting Bias	High	All measured outcomes outlined in the methods are reported in the results in a way that would allow for detailed extraction (effect estimates and 95% confidence intervals).
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Covariates included child sex, child race/ethnicity, median household income at baseline, duration of breastfeeding, gestational age, maternal and paternal allergy and asthma, birth weight was retrieved from medical records, and prenatal exposure to smoking. No information is provided on why these covariates were chosen, perhaps based on a priori knowledge.
	Metric 10:	Covariate Characterization	High	The majority of covariates were measured using questionnaires from the HOME study. Birth weight was retrieved from medical records, and prenatal smoking was measured using serum cotinine at ages 16 and 26 weeks of gestation.
	Metric 11:	Co-exposure Counfounding	Medium	The study measured and analyzed other co-exposures, including TCIPP, TDCIPP, TPHP, EH-TBB, and BEH-TEBP. While these were not adjusted for in analyses of TCEP, correlation coefficients were presented across all exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The prospective birth cohort study design is an appropriate design to assess the impact of prenatal TCEP exposure on respiratory outcomes. The use of generalized estimating equation analysis is appropriate to account for repeated measures.
	Metric 13:	Statistical Power	Medium	The number of participants in the analysis (n=342) is likely large enough to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficiently detailed so that the results can be reproduced.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Parent compound
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	High	The association between exposure and outcome was primarily measured using generalized estimating equations. The study chose a binomial distribution with a log link function for the models due to the repeated measures of the binary respiratory outcomes. Further, "unstructured working matrix and robust variance estimators were specified to estimate the beta coefficients exponentiated to obtain the relative risks and their corresponding 95% confidence intervals." Lung function parameters only had a single measurement and were thus assessed using linear regression modeling. Sensitivity analyses to assess the potential for effect modification by sex.Concentrations of TCEP were log-transformed to account for non-normality. TCEP was measured in 78.8% of all dust samples, but it is not specified how values below the LOD were handled. Analyses were presented separately for dust concentrations (ug/g dust) and dust loadings (mg/m^3).

Additional Comments: This prospective birth cohort study assessed the association between prenatal measurements of TCEP in dust samples and BCEP in maternal urine to repeated measures of respiratory outcomes among children. In general there is a limited risk of bias in the study due to well-conducted exposure assessment and outcome assessments. Significant positive associations were reported for TCEP and respiratory outcomes - however, these associations tended to be inconsistent across methods of exposure measurement (dust concentrations, dust loadings, urinary BCEP).

Overall Quality Determination High

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational weight gain among pregnant women; infant gestational age at delivery; infant anthropometric measurements at birth and 6 weeks postpartum, including birth weight and length, head and abdominal circumference, and four body composition (iliac, subscapular, triceps, and thigh skinfold thickness).
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Overall, all key elements of study design and participation were reported. In this pilot study, 62 women and infants were recruited. Inclusion and exclusion criteria as well as participant selection were fully described.
	Metric 2:	Attrition	High	A minimal number of subjects (6/62) were excluded from further analysis due to withdrawal, miscarriage, and lost to follow up. Exposure and outcome measurements were complete among pregnant women and infants. The exclusion or loss of follow up are not likely to introduce bias since the participated population is likely to represent the general eligible population.
	Metric 3:	Comparison Group	High	Differences in baseline characteristics were reported and adequately considered and adjusted in the statistical analysis. Identified effect modification by infant sex were considered in further analysis. Participants were similar since they were recruited from the same setting using same inclusion criteria.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure levels of the target chemical were derived from metabolite concentrations in urine samples. Collection and quantification methods were fully described.
	Metric 5:	Exposure Levels	Medium	Continuous individual exposure levels measured from OPE metabolite concentrations in pooled urine samples were used.
	Metric 6:	Temporality	Medium	The authors reported that OPE metabolites in urine samples showed good reproducibility and good intraclass correlation. They used pooled urine samples collected at gestation weeks 12, 28 and 35 to represent the exposure window. OPE chemicals have relatively short half lives and very low bioaccumulation rate, so the OPE chemicals are commonly used to indicate persistent exposure. It is not clear whether the exposures fell within relevant exposure windows for the outcomes of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Gestational weight gain (GWG) among nine pregnant women were substituted because of missing information. Newborn infants' anthropometry outcomes were measured twice by staff. A third measurement was applied if differences were out of pre-specified range. The same measurements were applied at 6 weeks postpartum. Significant measurement error is not likely to be present but not using the gold-standard.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational weight gain among pregnant women; infant gestational age at delivery; infant anthropometric measurements at birth and 6 weeks postpartum, including birth weight and length, head and abdominal circumference, and four body composition (iliac, subscapular, triceps, and thigh skinfold thickness).
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	High	All of the measured outcomes were described in detail. Effect measurements with 95% CI and medians with interquartile ranges were reported. Continuous exposure levels were used and each analysis was tabulated or graphed for data extraction.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariates including maternal age at delivery, income, pre-pregnancy BMI, parity and infant sex were appropriately adjusted in the linear regression model and mixed effect model.
	Metric 10:	Covariate Characterization	Medium	Covariate information was collected through a questionnaire at enrollment and medical records. Medical records are a well-established and reliable source, while the questionnaire is less-established. There is little to no concern about validity or confounding.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposure of 2 other OPE metabolites in urine samples were evaluated and adjusted in this study. Even though the authors mentioned residual confounding by unmeasured co-exposure may be present, there is no direct evidence that it would introduce significant bias to the effect.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design and analytical models used were appropriate to analyze the relationship between exposure and outcomes. Up to 3 urine metabolite measurements were used to represent the exposure level in the pregnancy window. Linear regression models were used for the continuous variables.
	Metric 13:	Statistical Power	Low	The authors reported that this pilot study with a smaller sample size (56 maternal-infant pairs) may not have sufficient statistical power to detect some effects. However, the power was sufficient to detect an association between OPE exposure and infant anthropometry by sex, which was also reported by Hoffman et al 2018.
	Metric 14:	Reproducibility of Analyses	Medium	The analysis methods, model selection, and data processing methods were reported and sufficient to understand and reproduce.
	Metric 15:	Statistical Analysis	High	Model assumptions were met and the method was transparent. Variable were appropriately transformed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	BCEP is a metabolite of TCEP and used as a biomarker of TCEP exposure. There might be other parent compounds but TCEP is one of the most common OPE detected.
	Metric 17:	Effect Biomarker	N/A	Not applicable - no biomarker of effect.
	Metric 18:	Method Sensitivity	Medium	Analytical methods were fully described and appropriate. The LODs and detection frequency were reported in the supplemental table S1.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational weight gain among pregnant women; infant gestational age at delivery; infant anthropometric measurements at birth and 6 weeks postpartum, including birth weight and length, head and abdominal circumference, and four body composition (iliac, subscapular, triceps, and thigh skinfold thickness).
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	High	Sample storage and shipping condition was reported and no reported loss.
	Metric 20:	Sample Contamination	High	There is no direct evidence to show the samples had contamination concerns. The analytical methods were described and the quality assurance used were within the lab limits, according to the authors.
	Metric 21:	Method Requirements	High	Target analytes were separated on a ultra-high-performanceliquid chromatography system and quantified using mass spectrometry.
	Metric 22:	Matrix Adjustment	N/A	The matrix adjustment information is not reported.

Additional Comments: Overall, this is a high-quality pilot study to evaluate the association between gestational exposure to target chemicals and reported health outcomes. The models and analytical methods applied were clear, fully described and appropriate. Strengths and limitations were discussed and not likely to introduce significant bias to the study.

Overall Quality Determination High

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Infant feeding behaviors including general appetite, enjoyment of food, food responsiveness, slowness in eating, and satiety responsiveness.
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Overall, all key elements of study design and participation were reported. In this pilot study, 62 women were recruited. Inclusion and exclusion criteria as well as participant selection were fully described.
	Metric 2:	Attrition	High	A minimal number of subjects (6/62) were excluded from further analysis due to withdrawal, miscarriage, and loss to follow up. Exposure and outcome measurements were complete among pregnant women and infants. The exclusion or loss to follow up are not likely to introduce bias since the participated population is likely to represent the general eligible population.
	Metric 3:	Comparison Group	High	Differences in baseline characteristics were reported and adequately considered and adjusted in statistical analysis. Identified effect modification by infant sex was considered in further analysis. Participants were similar since they were recruited from the same setting using same inclusion criteria.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure levels of target chemical were derived from metabolite concentrations in urine samples. Collection and quantification methods were fully described.
	Metric 5:	Exposure Levels	Medium	Continuous individual exposure levels were measured from OPE metabolite concentrations in pooled urine samples.
	Metric 6:	Temporality	Medium	The authors reported that OPE metabolites in urine samples showed good reproducibility and good intraclass correlation. They used pooled urine samples collected at gestation weeks 12, 28 and 35 to represent the exposure window. OPE chemicals have relatively short half lives and very low bioaccumulation rates, so the OPE chemicals are commonly used to indicate persistent exposure. It is not clear whether the exposures fell within relevant exposure windows for the outcomes of interests.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Infant feeding behaviors were evaluated by Baby Eating Behavior Questionnaire (BEBQ) completed by mothers. The validation is not reported but there is no direct evidence that the method has poor validity or significant misclassification.
	Metric 8:	Reporting Bias	High	All of the measured outcomes were described in detail. Effect measurements with 95% CI and medians with interquartile range reported. Continuous exposure levels were used and each analysis was tabulated or graphed for data extraction.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Infant feeding behaviors including general appetite, enjoyment of food, food responsiveness, slowness in eating, and satiety responsiveness.
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Covariates including maternal age at delivery, income, pre-pregnancy BMI, parity and infant sex were appropriately adjusted in the linear regression model and mixed effect model.
	Metric 10:	Covariate Characterization	Medium	Covariate information was collected through a questionnaire at enrollment and medical records. Medical records are a well-established and reliable source, while the questionnaire is less-established. There is little to no concern about validity or confounding.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposure of other 2 OPE metabolites in urine samples were evaluated and adjusted in this study. Even though the authors mentioned that residual confounding by unmeasured co-exposure may be present, there is no direct evidence that would introduce significant bias to the effect.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design and analytical models used were appropriate to catch the relationship between exposure and outcomes. Up to 3 urine metabolite measurements were applied to represent the exposure level during pregnancy. Linear regression models were used for the continuous variables.
	Metric 13:	Statistical Power	Low	The authors reported that this pilot study with a smaller sample size (56 maternal-infant pairs) may not have sufficient statistical power to detect some effects. However, the power was sufficient to detect an association between OPE exposure and infant anthropometry by sex, which was also reported by Hoffman et al 2018.
	Metric 14:	Reproducibility of Analyses	Medium	The description of analysis methods, model selection, and data processing methods were reported and sufficient to understand and reproduce.
	Metric 15:	Statistical Analysis	High	Model assumptions were met and the method was transparent. Variable were appropriately transformed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	BCEP is a metabolite of TCEP and used as a biomarker of TCEP exposure. There might be other parent compounds but TCEP is one of the most common OPE detected.
	Metric 17:	Effect Biomarker	N/A	Not applicable - no biomarker of effect.
	Metric 18:	Method Sensitivity	Medium	Analytical methods were fully described and appropriate. The LODs and detection frequency were reported in the supplemental table S1.
	Metric 19:	Biomarker Stability	High	Sample storage and shipping condition was reported and no reported loss.
	Metric 20:	Sample Contamination	High	There is no direct evidence to show the samples had contamination concerns. The analytical methods were described and the quality assurance used were within the lab limits, according to the authors.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Crawford, K. A., Hawley, N., Calafat, A. M., Jayatilaka, N. K., Froehlich, R. J., Has, P., Gallagher, L. G., Savitz, D. A., Braun, J. M., Werner, E. F., Romano, M. E. (2020). Maternal urinary concentrations of organophosphate ester metabolites: associations with gestational weight gain, early life anthropometry, and infant eating behaviors among mothers-infant pairs in Rhode Island. Environmental Health: A Global Access Science Source 19(1):97.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Infant feeding behaviors including general appetite, enjoyment of food, food responsiveness, slowness in eating, and satiety responsiveness.
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	7274557

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Target analytes were separated on a ultra-high-performanceliquid chromatography system and quantified using mass spectrometry.
	Metric 22:	Matrix Adjustment	N/A	The matrix adjustment information is not reported.

Additional Comments: Overall, this is a high-quality pilot study to evaluate the association between gestational exposure to target chemicals and reported health outcomes. Despite the small pilot-scale sample size, the samples are a good represent of eligible general population. The models and analytical methods applied were clear, fully described and appropriate. Strengths and limitations were discussed and not likely to introduce significant bias to the study.

Overall Quality Determination High

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants in this prospective birth cohort study were pulled from the Health Outcomes and Measures of the Environment (HOME) study. Pregnant women were recruited from March 2003-January 2006 in the greater Cincinnati, Ohio metropolitan area. Details of the HOME study are available in Braun et al. 2017 (HEROID: 6749104). In Braun et al. 2017 inclusion criteria and recruitment are extensively detailed and there is no evidence to suggest selection bias occurred in the original creation of the HOME study401 mother-child pairs were included in the original HOME study. Inclusion criteria for the analysis in this study were: having data on dust and urinary organophoshate esters and replacement brominated flame retardants, having data on covariates, and data on respiratory outcomes through follow-up until 5 years. 342 mother-child pairs were included in the present study.In general, there is no significant evidence for selection bias. However, no data is presented comparing those included in the study to participants who were excluded after originally participating in the HOME study.
	Metric 2:	Attrition	Medium	There were n=59 pregnant mother-child pair participants out of 401 who were included in the original HOME study but not the present analysis; all exclusion reasons were related to missing data, but exact breakdowns are not provided (ex; how many were excluded due to missing covariate data vs. missing exposure or outcome data). However, there is no reason to suggest that exclusion of those participants was inappropriate.
	Metric 3:	Comparison Group	High	Participants were all recruited from the same eligible population of pregnant women in the Greater Cincinnati, Ohio metropolitan area. The study presents a table of demographic characteristics split by exposure to all relevant exposures, including TCEP. Significant differences were reported by race/ethnicity and serum cotinine (significant differences were reported additionally for serum cotinine when assessing BCEP). All these factors were adjusted for in statistical analysis.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Exposure to TCEP was measured via the metabolite of BCEP. Urine was sampled at 16 and 26 weeks of gestation and within 48 hours of delivery. BCEP was quantified in spot urine samples using high-performance liquid chromatography mass spectrometry. The multitude of measurements allows for consistency across samples and a higher certainty in exposure classification.
	Metric 5:	Exposure Levels	Medium	Exposure ranges were likely large enough to allow an exposure-response estimate. For example, BCEP in urine average (5th-95th percentile): 0.13 - 6.98 ug/L.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	High	Exposure was measured at 16 weeks gestation, 26 weeks gestation, and at delivery. Temporality is thus established, and the introduction highlights the importance of developmental lung formation in predicting respiratory outcomes.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Respiratory outcomes included wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), and peak expiratory flow (PEF). Wheeze, respiratory infections, and hay fever or allergies were measured busting standardized questions, based on NHANES questionnaires. Questionnaires were provided every 6 months until the age of 5 years old. FEV1 and PEF were measured at age 5 years old using a portable spirometer, and measures were reported to follow the American Thoracic Society criteria and conducted by trained research assistants. In general, there is no significant evidence of outcome misclassification.
	Metric 8:	Reporting Bias	High	All measured outcomes outlined in the methods are reported in the results in a way that would allow for detailed extraction (effect estimates and 95% confidence intervals).
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Covariates included child sex, child race/ethnicity, median household income at baseline, duration of breastfeeding, gestational age, maternal and paternal allergy and asthma, birth weight was retrieved from medical records, and prenatal exposure to smoking. No information is provided on why these covariates were chosen.
	Metric 10:	Covariate Characterization	High	The majority of covariates were measured using questionnaires from the HOME study. Birth weight was retrieved from medical records, and prenatal smoking was measured using serum cotinine at ages 16 and 26 weeks of gestation.
	Metric 11:	Co-exposure Counfounding	Medium	The study measured and analyzed other co-exposures, including TCIPP, TDCIPP, TPHP, EH-TBB, and BEH-TEBP. While these were not adjusted for in analyses of TCEP, correlation coefficients were presented across all exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The prospective birth cohort study design is an appropriate design to assess the impact of prenatal TCEP exposure on respiratory outcomes. The use of generalized estimating equation analysis is appropriate to account for repeated measures.
	Metric 13:	Statistical Power	Medium	The number of participants in the analysis (n=342) is likely large enough to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficiently detailed so that the results can be reproduced.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mendy, A., Percy, Z., Braun, J. M., Lanphear, B., Guardia, La, M. J., Hale, R. C., Yolton, K., Chen, A. (2024). Prenatal exposure to replacement flame retardants and organophosphate esters and childhood adverse respiratory outcomes. Environmental Research 240(Pt 2):117523.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Wheeze, respiratory infections, hay fever or allergies, forced expiratory volume in 1s (FEV1), peak expiratory flow (PEF)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11364495

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	The association between exposure and outcome was primarily measured using generalized estimating equations. The study chose a binomial distribution with a log link function for the models due to the repeated measures of the binary respiratory outcomes. Further, "unstructured working matrix and robust variance estimators were specified to estimate the beta coefficients exponentiated to obtain the relative risks and their corresponding 95% confidence intervals." Lung function parameters only had a single measurement and were thus assessed using linear regression modeling. Sensitivity analyses to assess the potential for effect modification by sex.Concentrations of BCEP were log-transformed to account for non-normality. BCEP was detected in roughly 86% of samples across all three measured time periods, and samples below the LOD were replaced with LOD/sqrt(2). Analyses were presented separately for the urine sampling period.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP is a highly sensitive and specific biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	Effect biomarkers were not assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD is stated to be 0.10 ug/L, which is likely low enough to detect BCEP in a sufficient percentage of samples.
	Metric 19:	Biomarker Stability	Medium	Urine samples were frozen at -20 degrees Celsius until analysis. No stability data is reported.
	Metric 20:	Sample Contamination	Medium	There is no information provided on sample contamination, but there is no reason to suspect contamination occurred.
	Metric 21:	Method Requirements	High	Samples were analyzed using high performance liquid chromatography with mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Samples were adjusted for urinary dilution by standardizing for specific gravity. Results are only presented as adjusted.

Additional Comments: This prospective birth cohort study assessed the association between prenatal measurements of TCEP in dust samples and BCEP in maternal urine to repeated measures of respiratory outcomes among children. In general there is a limited risk of bias in the study due to well-conducted exposure assessment and outcome assessments. Significant positive associations were reported for TCEP and respiratory outcomes - however, these associations tended to be inconsistent across methods of exposure measurement (dust concentrations, dust loadings, urinary BCEP).

Overall Quality Determination High

Study Citation:	Percy, Z., Vuong, A. M., Xu, Y., Xie, C., Ospina, M., Calafat, A. M., Lanphear, B. P., Braun, J. M., Cecil, K. M., Dietrich, K. N., Chen, A., Yolton, K. (2021). Prenatal exposure to a mixture of organophosphate esters and intelligence among 8-year-old children of the HOME Study. NeuroToxicology 87:149-155.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Weschler Intelligence Scale for Children-IV scores (FSIQ, Verbal Comprehension, Perceptual Reasoning, Working Memory, Processing Speed)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	10081087

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	The HOME study recruited pregnant women from Cincinnati, Ohio from 2003 to 2006. Inclusion criteria were women 18 years of age or older, at 13-19 weeks of gestation, and living in a home built before 1978. Women were excluded if they were taking medication for thyroid disorders or seizures, HIV positive, had bipolar disorder, schizophrenia, diabetes, active cancer, not fluent in English, or planning to move outside the Cincinnati area. 5184 women were approached about the study, 1263 met the eligibility criteria and 468 women agreed to participate (37% participation rate). Entry into this specific analysis of HOME study participants also required that women deliver a live, singleton infant, have at least one urinary OPE measurement during pregnancy, and have a child intelligence measurement at 8 or 12 years of age. In total, 233 participants participated in the final analysis.There was no comparison between women who joined and those who did not join the study and limited the methods of study advertisement and recruitment. However, there is no evidence of selection bias.
	Metric 2:	Attrition	Medium	The participation rate after follow up at 8-12 years of age was 49.8% (n=233). Sociodemographic characteristics of the remaining participants and the original participants were reported to be similar. There was no evidence that missingness was related to both exposure and outcome.
	Metric 3:	Comparison Group	Medium	Demographic characteristics are not presented in association with exposure or outcome levels, and thus there is no direct evidence that groups were similar. However, there is no reason to suspect that groups were wildly different, as key demographic characteristics were adjusted for in statistical analysis. Recruitment methods and inclusion/exclusion do not appear to have been performed differently in subsets of participants.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	TCEP was measured via the urinary metabolite BCEP in spot urine samples collected at 3 times over the course of pregnancy (around 16 weeks' gestation, 26 weeks' gestation, and within 48 hours of delivery). BCEP was quantified using high performance liquid chromatography with tandem mass spectrometry. The use of multiple measurement time points increases the confidence in reported exposure levels.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Vuong, A. M., Xu, Y., Xie, C., Ospina, M., Calafat, A. M., Lanphear, B. P., Braun, J. M., Cecil, K. M., Dietrich, K. N., Chen, A., Yolton, K. (2021). Prenatal exposure to a mixture of organophosphate esters and intelligence among 8-year-old children of the HOME Study. NeuroToxicology 87:149-155.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Weschler Intelligence Scale for Children-IV scores (FSIQ, Verbal Comprehension, Perceptual Reasoning, Working Memory, Processing Speed)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	10081087

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	The median value for samples collected at 16 weeks was 0.76 ng/mL (25th - 75th percentile: 0.45 - 1.34 ng/mL). The median value for samples collected at 26 weeks was 0.65 ng/mL (25th - 75th percentile: 0.34 - 1.31 ng/mL). The median value for samples collected at delivery was 0.62 ng/mL (25th - 75th percentile: 0.35 - 1.27 ng/mL). Generally, the levels of exposure at all three time points likely have sufficient variation to detect an effect. Concentrations of BCEP were modeled continuously in statistical models.
	Metric 6:	Temporality	Medium	Exposure measures were taken three times before the outcome during pregnancy, so temporality was established and there is some degree of certainty that exposure levels were consistent across pregnancy. However, due to the short half-life of BCEP and a long follow-up time, it's unclear whether the exposure was consistent across the first 8 or 12 years of life, and it is unclear whether or not gestation represents the only etiologically relevant time period for neurodevelopmental effects.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The outcome was assessed by using Wechsler Intelligence Scale for Children-IV (WISC-IV) for children at 8 years old. A small number of children (n=12) were assessed at 12 years old using the same questionnaire. The tests were administered by research assistants trained by an experienced and reliable gold standard examiner and blinded to all exposure variables. Training exercises for examiners are described in detail. Examiners were re-assessed every six months to ensure accuracy in assessment.
	Metric 8:	Reporting Bias	High	All analyses mentioned in the methods are reported in the results. Effect estimates were reported as regression coefficient with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Covariates were identified through a directed acyclic graph and kept minimum for analysis. The covariates were income, maternal BMI, maternal IQ as measured by WASI, and maternal race.
	Metric 10:	Covariate Characterization	Medium	Covariate information was measured using standardized questionnaire. Maternal intelligence was assessed with WASI and examination methods are extensively detailed.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposures included other OPEs (BDCIPP, DPHP, DNBP) and 10 PBDE congeners (BDE-17, -28, -66, -85, -99, -100, -153, -154, -183). Detailed methods are provided for the quantification of each co-exposure. Exposure to PBDEs was adjusted for in statistical analyses, while co-exposure to other OPEs was addressed via BKMR modeling.
Domain 5: Analysis				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Vuong, A. M., Xu, Y., Xie, C., Ospina, M., Calafat, A. M., Lanphear, B. P., Braun, J. M., Cecil, K. M., Dietrich, K. N., Chen, A., Yolton, K. (2021). Prenatal exposure to a mixture of organophosphate esters and intelligence among 8-year-old children of the HOME Study. NeuroToxicology 87:149-155.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Weschler Intelligence Scale for Children-IV scores (FSIQ, Verbal Comprehension, Perceptual Reasoning, Working Memory, Processing Speed)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	10081087

Domain	Metric		Rating	Comments
	Metric 12:	Study Design and Methods	High	The study design was a prospective birth cohort with a long follow-up time to evaluate the association between maternal urinary BCEP exposure and child intelligence. The use of repeated measures analysis is also appropriate to assess changes in exposure levels of BCEP across pregnancy.
	Metric 13:	Statistical Power	Medium	The sample size (N=233 mother-child dyads) is likely large enough to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The methods are described in sufficient detail so that they would able to be reproduced given access to the analytic data.
	Metric 15:	Statistical Analysis	High	The association between prenatal concentration of BCEP and child intelligence was assessed using generalized estimating equations to account for the three different measurements during the pregnancy. BCEP concentrations were ln-transformed to normalize the distribution. As BCEP had more than 10% of samples with concentrations below the limit of detection, a multiple imputation approach was used to impute left-censored data via Markov Chain Monte Carlo algorithm to yield 10 imputed data sets. Each data set was used for modeling and the results were pooled using Rubin's rules. The results were presented as regression coefficient with 95% confidence interval.Sensitivity analysis was conducted to include urinary specific gravity as a covariate and to standardize OPE concentrations by specific gravity.In general, the statistical analysis method was appropriate, and model assumptions were met.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	The biomarker BCEP is derived from the parent compound TCEP and is appropriate to use as a proxy for exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	0
	Metric 18:	Method Sensitivity	Medium	The limit of detection was 0.1 ng/mL and is likely low enough to detect chemicals in a sufficient percentage of samples.
	Metric 19:	Biomarker Stability	Medium	Samples were collected and kept frozen at -20 degrees Celsius for storage. No stability data was reported but there was no evidence of sample loss
	Metric 20:	Sample Contamination	Medium	There is no specific information about contamination, but quality control procedures are detailed.
	Metric 21:	Method Requirements	High	BCEP was quantified via high-performance liquid chromatography with isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	BCEP concentrations were only presented as adjusted for specific gravity.

Additional Comments: The HOME prospective birth cohort had a good design to evaluate any associations between prenatal exposure to BCEP and child intelligence at ages 8 or 12. There were no significant sources of potential bias, and the exposure assessment was well-conducted with urine sampled three times during pregnancy to ensure consistent exposure during the gestational period. BCEP was reported to be associated with a slight increase in child full-scale IQ.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Vuong, A. M., Xu, Y., Xie, C., Ospina, M., Calafat, A. M., Lanphear, B. P., Braun, J. M., Cecil, K. M., Dietrich, K. N., Chen, A., Yolton, K. (2021). Prenatal exposure to a mixture of organophosphate esters and intelligence among 8-year-old children of the HOME Study. NeuroToxicology 87:149-155.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Weschler Intelligence Scale for Children-IV scores (FSIQ, Verbal Comprehension, Perceptual Reasoning, Working Memory, Processing Speed)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	10081087

Domain	Metric	Rating	Comments
Overall Quality Determination		**High**	

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Study used data from NHANES 2013-2014. Cohort is designed to be representative sample of general US population. Analyzed random selection of approximately 1/3 of participants with stored spot urine samples. Inclusion/exclusion criteria were appropriate for the outcome assessment: excluded pregnant women, those with missing eGFR, ACR or BMI data. Low risk of selection bias.
	Metric 2:	Attrition	High	Of the 2666 stored urine samples chosen to measure OPE metabolites, 1578 were used in the analysis for this study. The exclusion of participants was appropriately addressed based on age, current pregnancy (1,660). BMI, ACR, or eGFR data was missing for only 82 (and subsequently excluded). Exposure and outcome data largely complete.
	Metric 3:	Comparison Group	High	Baseline characteristics of groups assessed and adjusted for in statistical modelling, including sex, age, race/ethnicity, BMI, smoking status, poverty income ratio, physical and current CKD. Large established cohort with documented setting, methods of selection. (Documented previously) Appropriate inclusion/exclusion criteria for study reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	The investigators used an established method for analyzing the exposure (biomarker): the analytes were separated using reverse-phase high performance liquid chromatography. The analytes were measured using isotope dilution-electroscopy ionization tandem mass spectrometry. The study also measured urine creatinine to account for urinary dilution and concentration of the biomarkers.
	Metric 5:	Exposure Levels	Medium	The authors reported LOD for the analytes which were detected in >75% of samples, and also the detection frequency, geometric mean and 25, 50 and 75 percentile. The distributions of metabolites were reported for unadjusted, adjusted using a traditional creatinine adjustment and using a novel creatinine adjustment.
	Metric 6:	Temporality	Low	Cross-sectional study precludes temporal causality. In addition, this study utilized spot urine sample. The metabolites have short half-life and thus it is difficult to assess long-term exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	This study used eGFR and ACR parameters to classify participants as chronic kidney disease (CKD) patients; the parameters previously established using an extended definition of CKD. eGFR was calculated using an appropriate formula. The ACR was calculated as the urinary albumin/creatinine ratio.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	High	The authors report all of the study's measured outcomes as described in the methods. N is reported for each analysis, and 95% CI reported for associations.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	This study accounted for several potential confounders and the analyses adjusted for: sex, age, race/ethnicity, BMI, poverty income ratio, smoking history, and physical activity.
	Metric 10:	Covariate Characterization	Medium	NHANES uses a validated survey instrument for obtaining potential confounder data. The urinary creatinine (and thus urine dilution) was determined using established laboratory methods. Use of creatine to establish urine dilution does have limitations and warrants further examination. This study used two different methods, one traditional and another "novel" approach. These different methods revealed changes in the association between kidney function and metabolite concentration. The authors state "Since kidney function can directly influence urine hydration or creatinine excretion, however, the use of creatinine-adjustment for urine dilution may induce a collider stratification biasleading to possible confounding".
	Metric 11:	Co-exposure Counfounding	Medium	Correlation co-efficients were calculated for the OPE metabolites. Co-exposures to pollutants (biomarkers -metabolites) were assessed in the analysis using a multi-pollutant model.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. Limited simply by the design. However, the statistical methods used for this study were appropriate for the data.
	Metric 13:	Statistical Power	Medium	The sample size (n = 1578) was adequate to detect an effect, power of study not reported.
	Metric 14:	Reproducibility of Analyses	Medium	The authors clearly describe the analytic methods and models used for the analyses. They could be reproduced.
	Metric 15:	Statistical Analysis	High	The models and methods used to analyze the associations between exposure and outcomes were appropriate, and the inclusion/exclusion of variables was clearly stated. Model assumptions were met.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Low	The authors note that the metabolism of OPE is not completely understood. OPE is metabolized into diester metabolites and excreted though urine; this is what was measured in this study. No discussion of validation or assessment of accuracy and precision.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
	Metric 17:	Effect Biomarker	Low	Little information available on OPE toxic effect on kidney outcomes. There is some evidence of potential health effects on humans and animals (disruption of thyroid and sex hormones, development and behavior, and allergic response) but mechanism of action is not well understood.
	Metric 18:	Method Sensitivity	Medium	The LOD are low enough to detect the metabolites in a sufficient percentage of samples. The analytic methods are described in detail and use established methods. The LOD and % are reported.
	Metric 19:	Biomarker Stability	Low	Storage history and stability data for analytes is not reported. OPE metabolites have a short half life, but this would be reflected in the cross section of samples taken and not a reflection of the stability in properly stored samples.
	Metric 20:	Sample Contamination	Low	There is no information regarding storage or potential contamination and subsequent steps to correct issues. However, there is no reason to suspect there are any issues with this large and well-established cohort's sampling methods.
	Metric 21:	Method Requirements	High	This study used LC-MS/MS for mesuring the metabolites.
	Metric 22:	Matrix Adjustment	High	This study reported both adjusted and unadjusted matrix concentrations (used 2 different methods of urinary creatinine adjustment) and this is supported and discussed at length due to the importance of urine dilution in studies of kidney function.

Additional Comments: This is a large cross sectional study using NHANES 2013-2014 data to examine the associations between urinary OPE's and CKD. This study used appropriate participant selection, analytic methods and examined important potential confounders. However, the cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. There is limited ability to determine causal effect of exposure on the outcome. The mechanisms of OPE and metabolites' effect on CKD need further investigating.

Overall Quality Determination Medium

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Study used data from NHANES 2013-2014. Cohort is designed to be representative sample of general US population. Analyzed random selection of approximately 1/3 of participants with stored spot urine samples. Inclusion/exclusion criteria were appropriate for the outcome assessment: excluded pregnant women, those with missing eGFR, ACR or BMI data. Low risk of selection bias.
	Metric 2:	Attrition	High	Of the 2666 stored urine samples chosen to measure OPE metabolites, 1578 were used in the analysis for this study. The exclusion of participants was appropriately addressed based on age, current pregnancy (1,660). BMI, ACR, or eGFR data was missing for only 82 (and subsequently excluded). Exposure and outcome data largely complete.
	Metric 3:	Comparison Group	High	Baseline characteristics of groups assessed and adjusted for in statistical modelling, including sex, age, race/ethnicity, BMI, smoking status, poverty income ratio, physical and current CKD. Large established cohort with documented setting, methods of selection. (Documented previously) Appropriate inclusion/exclusion criteria for study reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	The investigators used an established method for analyzing the exposure (biomarker): the analytes were separated using reverse-phase high performance liquid chromatography. The analytes were measured using isotope dilution-electroscopy ionization tandem mass spectrometry. The study also measured urine creatinine to account for urinary dilution and concentration of the biomarkers.
	Metric 5:	Exposure Levels	Medium	The authors reported LOD for the analytes which were detected in >75% of samples, and also the detection frequency, geometric mean and 25, 50 and 75 percentile. The distributions of metabolites were reported for unadjusted, adjusted using a traditional creatinine adjustment and using a novel creatinine adjustment.
	Metric 6:	Temporality	Low	Cross-sectional study precludes temporal causality. In addition, this study utilized spot urine sample. The metabolites have short half-life and thus it is difficult to assess long-term exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	This study used eGFR and ACR parameters to classify participants as chronic kidney disease (CKD) patients; the parameters previously established using an extended definition of CKD. eGFR was calculated using an appropriate formula. The ACR was calculated as the urinary albumin/creatinine ratio.
	Metric 8:	Reporting Bias	High	The authors report all of the study's measured outcomes as described in the methods. N is reported for each analysis, and 95% CI reported for associations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	This study accounted for several potential confounders and the analyses adjusted for: sex, age, race/ethnicity, BMI, poverty income ratio, smoking history, and physical activity.
	Metric 10:	Covariate Characterization	Medium	NHANES uses a validated survey instrument for obtaining potential confounder data. The urinary creatinine (and thus urine dilution) was determined using established laboratory methods. Use of creatine to establish urine dilution does have limitations and warrants further examination. This study used two different methods, one traditional and another "novel" approach. These different methods revealed changes in the association between kidney function and metabolite concentration. The authors state "Since kidney function can directly influence urine hydration or creatinine excretion, however, the use of creatinine-adjustment for urine dilution may induce a collider stratification biasleading to possible confounding".
	Metric 11:	Co-exposure Counfounding	Medium	Correlation co-efficients were calculated for the OPE metabolites. Co-exposures to pollutants (biomarkers -metabolites) were assessed in the analysis using a multi-pollutant model.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. Limited simply by the design. However, the statistical methods used for this study were appropriate for the data.
	Metric 13:	Statistical Power	Medium	The sample size (n = 1578) was adequate to detect an effect, power of study not reported.
	Metric 14:	Reproducibility of Analyses	Medium	The authors clearly describe the analytic methods and models used for the analyses. They could be reproduced.
	Metric 15:	Statistical Analysis	High	The models and methods used to analyze the associations between exposure and outcomes were appropriate, and the inclusion/exclusion of variables was clearly stated. Model assumptions were met.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Low	The authors note that the metabolism of OPE is not completely understood. OPE is metabolized into diester metabolites and excreted though urine; this is what was measured in this study. No discussion of validation or assessment of accuracy and precision.
	Metric 17:	Effect Biomarker	Low	Little information available on OPE toxic effect on kidney outcomes. There is some evidence of potential health effects on humans and animals (disruption of thyroid and sex hormones, development and behavior, and allergic response) but mechanism of action is not well understood.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
	Metric 18:	Method Sensitivity	Medium	The LOD are low enough to detect the metabolites in a sufficient percentage of samples. The analytic methods are described in detail and use established methods. The LOD and % are reported.
	Metric 19:	Biomarker Stability	Low	Storage history and stability data for analytes is not reported. OPE metabolites have a short half life, but this would be reflected in the cross section of samples taken and not a reflection of the stability in properly stored samples.
	Metric 20:	Sample Contamination	Low	There is no information regarding storage or potential contamination and subsequent steps to correct issues. However, there is no reason to suspect there are any issues with this large and well-established cohort's sampling methods.
	Metric 21:	Method Requirements	High	This study used LC-MS/MS for mesuring the metabolites.
	Metric 22:	Matrix Adjustment	High	This study reported both adjusted and unadjusted matrix concentrations (used 2 different methods of urinary creatinine adjustment) and this is supported and discussed at length due to the importance of urine dilution in studies of kidney function.

Additional Comments: This is a large cross sectional study using NHANES 2013-2014 data to examine the associations between urinary OPE's and CKD. This study used appropriate participant selection, analytic methods and examined important potential confounders. However, the cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. There is limited ability to determine causal effect of exposure on the outcome. The mechanisms of OPE and metabolites' effect on CKD need further investigating.

Overall Quality Determination Medium

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Study used data from NHANES 2013-2014. Cohort is designed to be representative sample of general US population. Analyzed random selection of approximately 1/3 of participants with stored spot urine samples. Inclusion/exclusion criteria were appropriate for the outcome assessment: excluded pregnant women, those with missing eGFR, ACR or BMI data. Low risk of selection bias.
	Metric 2:	Attrition	High	Of the 2666 stored urine samples chosen to measure OPE metabolites, 1578 were used in the analysis for this study. The exclusion of participants was appropriately addressed based on age, current pregnancy (1,660). BMI, ACR, or eGFR data was missing for only 82 (and subsequently excluded). Exposure and outcome data largely complete.
	Metric 3:	Comparison Group	High	Baseline characteristics of groups assessed and adjusted for in statistical modelling, including sex, age, race/ethnicity, BMI, smoking status, poverty income ratio, physical and current CKD. Large established cohort with documented setting, methods of selection. (Documented previously) Appropriate inclusion/exclusion criteria for study reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	The investigators used an established method for analyzing the exposure (biomarker): the analytes were separated using reverse-phase high performance liquid chromatography. The analytes were measured using isotope dilution-electroscopy ionization tandem mass spectrometry. The study also measured urine creatinine to account for urinary dilution and concentration of the biomarkers.
	Metric 5:	Exposure Levels	Medium	The authors reported LOD for the analytes which were detected in >75% of samples, and also the detection frequency, geometric mean and 25, 50 and 75 percentile. The distributions of metabolites were reported for unadjusted, adjusted using a traditional creatinine adjustment and using a novel creatinine adjustment.
	Metric 6:	Temporality	Low	Cross-sectional study precludes temporal causality. In addition, this study utilized spot urine sample. The metabolites have short half-life and thus it is difficult to assess long-term exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	This study used eGFR and ACR parameters to classify participants as chronic kidney disease (CKD) patients; the parameters previously established using an extended definition of CKD. eGFR was calculated using an appropriate formula. The ACR was calculated as the urinary albumin/creatinine ratio.
	Metric 8:	Reporting Bias	High	The authors report all of the study's measured outcomes as described in the methods. N is reported for each analysis, and 95% CI reported for associations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	This study accounted for several potential confounders and the analyses adjusted for: sex, age, race/ethnicity, BMI, poverty income ratio, smoking history, and physical activity.
	Metric 10:	Covariate Characterization	Medium	NHANES uses a validated survey instrument for obtaining potential confounder data. The urinary creatinine (and thus urine dilution) was determined using established laboratory methods. Use of creatine to establish urine dilution does have limitations and warrants further examination. This study used two different methods, one traditional and another "novel" approach. These different methods revealed changes in the association between kidney function and metabolite concentration. The authors state "Since kidney function can directly influence urine hydration or creatinine excretion, however, the use of creatinine-adjustment for urine dilution may induce a collider stratification biasleading to possible confounding".
	Metric 11:	Co-exposure Counfounding	Medium	Correlation co-efficients were calculated for the OPE metabolites. Co-exposures to pollutants (biomarkers -metabolites) were assessed in the analysis using a multi-pollutant model.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. Limited simply by the design. However, the statistical methods used for this study were appropriate for the data.
	Metric 13:	Statistical Power	Medium	The sample size (n = 1578) was adequate to detect an effect, power of study not reported.
	Metric 14:	Reproducibility of Analyses	Medium	The authors clearly describe the analytic methods and models used for the analyses. They could be reproduced.
	Metric 15:	Statistical Analysis	High	The models and methods used to analyze the associations between exposure and outcomes were appropriate, and the inclusion/exclusion of variables was clearly stated. Model assumptions were met.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Low	The authors note that the metabolism of OPE is not completely understood. OPE is metabolized into diester metabolites and excreted though urine; this is what was measured in this study. No discussion of validation or assessment of accuracy and precision.
	Metric 17:	Effect Biomarker	Low	Little information available on OPE toxic effect on kidney outcomes. There is some evidence of potential health effects on humans and animals (disruption of thyroid and sex hormones, development and behavior, and allergic response) but mechanism of action is not well understood.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Kang, H., Lee, J., Lee, J. P., Choi, K. (2019). Urinary metabolites of organophosphate esters (OPEs) are associated with chronic kidney disease in the general US population, NHANES 2013–2014. Environment International 131:5034-5034.
Health Outcome(s) and Reported Health Effect(s):	Renal/Kidney-estimated Glomerular Filtration Rate (eGFR)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: BCEP
Linked HERO ID(s):	No linked references.
HERO ID:	10078361

Domain	Metric		Rating	Comments
	Metric 18:	Method Sensitivity	Medium	The LOD are low enough to detect the metabolites in a sufficient percentage of samples. The analytic methods are described in detail and use established methods. The LOD and % are reported.
	Metric 19:	Biomarker Stability	Low	Storage history and stability data for analytes is not reported. OPE metabolites have a short half life, but this would be reflected in the cross section of samples taken and not a reflection of the stability in properly stored samples.
	Metric 20:	Sample Contamination	Low	There is no information regarding storage or potential contamination and subsequent steps to correct issues. However, there is no reason to suspect there are any issues with this large and well-established cohort's sampling methods.
	Metric 21:	Method Requirements	High	This study used LC-MS/MS for mesuring the metabolites.
	Metric 22:	Matrix Adjustment	High	This study reported both adjusted and unadjusted matrix concentrations (used 2 different methods of urinary creatinine adjustment) and this is supported and discussed at length due to the importance of urine dilution in studies of kidney function.

Additional Comments: This is a large cross sectional study using NHANES 2013-2014 data to examine the associations between urinary OPE's and CKD. This study used appropriate participant selection, analytic methods and examined important potential confounders. However, the cross-sectional design used spot urine levels to establish association with exposure and a chronic disease outcome. There is limited ability to determine causal effect of exposure on the outcome. The mechanisms of OPE and metabolites' effect on CKD need further investigating.

Overall Quality Determination Medium

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z. Z., Kannan, K., Robinson, M., Foley, H. B., Grubbs, B., Al-Marayati, L., Lerner, D., Lurvey, N., Aung, M. T., Habre, R., Dunton, G. F., Farzan, S. F., Breton, C. V., Bastain, T. M. (2023). Sex-specific effects of prenatal organophosphate ester (OPE) metabolite mixtures and adverse infant birth outcomes in the maternal and developmental risks from environmental and social stressors (MADRES) pregnancy cohort. Environmental Research 226:115703.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, gestational-age adjusted birth weight z-scores
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577122

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	This study used data from 421 mother-infant pairs in the ongoing MADRES (Maternal And Developmental Risks from Environmental and Social Stressors) cohort to analyze associations between maternal urinary concentrations of the TCEP metabolite bis(2-chloroethyl) phosphate (BCEP) and birth outcomes. This cohort has recruited predominately low-income Hispanic/Latino mother-child pairs from Los Angeles; participants in this study were recruited from 2015 and 2019. Eligible women were aged 18 years or older, had no significant disabilities or HIV, were recruited prior to 30 weeks' gestation, and had singleton pregnancies. Participation rates for the cohort were not described. However, there was no evidence to suggest that participation would have been associated with either TCEP exposure or birth outcomes. A subset of 421 participants with OPE metabolite concentrations, birth outcome data, and information on key covariates was created for this analysis and the study reported that this subset was similar to the full cohort on "key demographic characteristics" (supporting data shown in related HEROID 11581666, parent cohort N=774).
	Metric 2:	Attrition	Medium	The analysis sample included MADRES participants with available urinary organophosphate ester (OPE) metabolites. Urinary OPEs, including BCEP, were measured in all available third trimester spot urine samples collected from women recruited from November 2015 to October 2019. Of 426 participants, 5 were excluded due to missing information on key covariates. The authors stated that the subset of 421 was similar to the parent cohort (supporting data shown in related HEROID 11581666, parent cohort N=774). Overall, the analysis sample included more than half of the parent cohort, with minimal attrition among those with available exposure data. There was no evidence of any bias.
	Metric 3:	Comparison Group	Medium	Analyses compared birth outcome measures among participants with lower vs. higher BCEP in maternal urine during gestation. While descriptive data stratified by BCEP were not presented, there was no evidence to suggest the comparison group with lower BCEP differed from those with higher concentrations, and many potential demographic differences (ex: age, income, education) were considered as potential covariates.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z. Z., Kannan, K., Robinson, M., Foley, H. B., Grubbs, B., Al-Marayati, L., Lerner, D., Lurvey, N., Aung, M. T., Habre, R., Dunton, G. F., Farzan, S. F., Breton, C. V., Bastain, T. M. (2023). Sex-specific effects of prenatal organophosphate ester (OPE) metabolite mixtures and adverse infant birth outcomes in the maternal and developmental risks from environmental and social stressors (MADRES) pregnancy cohort. Environmental Research 226:115703.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, gestational-age adjusted birth weight z-scores
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577122

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	Prenatal TCEP exposure was estimated using BCEP concentrations from a single maternal urine sample collected at a mean of 31.5 weeks gestation, adjusted for dilution using specific gravity. BCEP was detected in 68.4% of samples. Values below the LOD of 0.02 ng/mL were imputed as LOD divided by the square root of 2. Use of a biomarker is a strength. A potential limitation is the use of a single spot urine sample, given that within-person variability in BCEP throughout pregnancy is unknown. Other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372).; excluding this factor may also introduce measurement error. Though some measurement error is likely, there was no evidence of important error or bias in exposure estimates.
	Metric 5:	Exposure Levels	Medium	BCEP was analyzed as a continuous variable. The distribution indicated adequate variability, with a median (IQR) of 0.53 (0.03, 1.62) ng/mL.
	Metric 6:	Temporality	Medium	BCEP concentrations were measured at a mean (SD) of 31.5 (2) weeks' gestation, preceding child behavior measures characterized at age 36 months. BCEP concentrations were not measured in early or mid-pregnancy. However, there was no evidence of specific vulnerable windows or other concerns for temporality. It is uncertain whether there may be potentially sensitive windows for adverse effects of TCEP on birth weight or gestational age
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcomes analyzed in this study were gestational age at birth, birth weight, and gestational-age adjusted birth weight z-scores. Gestational age was estimated using the following hierarchy: first trimester ultrasound (59.6%), second trimester ultrasound, and physician clinical estimate in medical records; the estimate for one subject used last menstrual period. 79% of participants had multiple gestational age estimates, with Spearman correlations reported as ranging from 0.52 to 0.91. Correlations of 0.52 suggest important variability in estimates. Though there was no further information on the extent of variation in gestational age estimates, more accurate estimates were prioritized in the selection hierarchy. 9.3% of participants were classified as preterm (<37 weeks' gestation). Birth weight was extracted from medical records and used to calculate sex-specific birthweight for gestational age z-scores using a nationally representative U.S. sample. Because unadjusted birth weight was not analyzed, any measurement error in gestational age is integrated. Despite concerns, there was no evidence of important errors in outcome characterization.
	Metric 8:	Reporting Bias	High	Results were presented or described for all analyses included as aims.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z. Z., Kannan, K., Robinson, M., Foley, H. B., Grubbs, B., Al-Marayati, L., Lerner, D., Lurvey, N., Aung, M. T., Habre, R., Dunton, G. F., Farzan, S. F., Breton, C. V., Bastain, T. M. (2023). Sex-specific effects of prenatal organophosphate ester (OPE) metabolite mixtures and adverse infant birth outcomes in the maternal and developmental risks from environmental and social stressors (MADRES) pregnancy cohort. Environmental Research 226:115703.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, gestational-age adjusted birth weight z-scores
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577122

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Covariates were selected based on prior literature and using a directed acyclic graph (DAG). All multivariate models adjusted for recruitment site, season of sample collection, gestational age at sample collection, maternal age, pre-pregnancy BMI, parity, race/ethnicity, maternal hypertensive disorders of pregnancy, income, and education). Infant sex was included except for models stratified by sex or analyzing sex-specific birthweight for gestational age; smoking during pregnancy was addressed by exclusion in sensitivity analysis due to its low prevalence. Gestational diabetes was also included as a covariate in sensitivity analyses. Although all analyses of birth weight included gestational age, which is a potential intermediate, the finding of no association between TCEP and gestational age mitigates concern of important bias. Gestational hypertension is also not established as a potential intermediate. While there are potential concerns related to overadjustment, there was no evidence of important bias or of residual confounding.
	Metric 10:	Covariate Characterization	Medium	Covariate data came from interviewer-administered questionnaire or medical records. Validation was not discussed explicitly, but there was no evidence to suggest important error or bias.
	Metric 11:	Co-exposure Counfounding	Medium	Correlations between BCEP and other OPEs (DPHP, DNBP, DIBP, BDCIPP, BBOEP, BCIPP, BMPP, BEHP, DPRP) measured in this study were weak (Spearman's r <0.25). Pair-wise interactions among OPE metabolites were examined. Co-exposure to other OPEs during pregnancy was also evaluated using Bayesian methods to analyze chemical mixtures. Though confounding by other exposures cannot be ruled out there was no evidence indicative of important co-exposure confounding.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	A prospective cohort design was used to analyze the relationship between prenatal TCEP exposure and birth outcomes, with exposure measured in the third trimester. This design was appropriate. Linear regression was appropriately used for analysis.
	Metric 13:	Statistical Power	Medium	The analysis sample included 421 participants with a mean (SD) gestational age of 39.1 (1.5) weeks, birthweight of 3303 (475) grams, and birthweight for gestational age z-score of -0.1 (1.0). There was variability in exposure. There was no indication of inadequate statistical power.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what was done and to replicate findings.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z. Z., Kannan, K., Robinson, M., Foley, H. B., Grubbs, B., Al-Marayati, L., Lerner, D., Lurvey, N., Aung, M. T., Habre, R., Dunton, G. F., Farzan, S. F., Breton, C. V., Bastain, T. M. (2023). Sex-specific effects of prenatal organophosphate ester (OPE) metabolite mixtures and adverse infant birth outcomes in the maternal and developmental risks from environmental and social stressors (MADRES) pregnancy cohort. Environmental Research 226:115703.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, gestational-age adjusted birth weight z-scores
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577122

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were presented for the sample as a whole. Urinary BCEP was described using percentiles, the geometric means, and the minimum and maximum. To satisfy model assumptions and evaluate potentially non-linear dose-response, primary models for individual exposures analyzed associations with natural-log transformed CBCL scores using either BCEP tertiles or using generalized additive models. Both crude and adjusted effect estimates were presented with 95% confidence intervals. Interaction terms and stratified models were used to assess effect modification by child sex. In addition, Bayesian kernel machine regression (BKMR) models were used to analyze OPE mixtures, and Bayesian semiparametric regression models were used to analyze synergistic or antagonistic interactions among pairs of chemicals. Both sets of models allowed for non-linear dose-response patterns. Results were presented graphically and using posterior inclusion probabilities with 95% credible intervals. Sensitivity analyses to assess robustness included excluding children whose mothers smoked during pregnancy, analyzing CBCL t-scores vs raw scores, and adjusting BKMR smoothness parameters. There was no evidence of deficiencies in the data analysis methods used.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Exposure was estimated using urinary concentrations of the major TCEP metabolite BCEP measured in a spot urine sample from the third trimester of pregnancy. Unmetabolized TCEP, detected in 37% of urine samples in another study (Hou et al. 2020 HEROID 10143372), was not measured. BCEP concentrations in a single spot urine sample characterize TCEP exposure in pregnancy with some measurement error, but there was no evidence of bias.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD was 0.02 ng/mL; 68.41% of samples had concentrations above detection limits.
	Metric 19:	Biomarker Stability	Medium	Aliquoted samples were stored at -80 degrees Celsius prior to shipment. Storage duration was not reported, but there was no evidence of degradation or inappropriate handling. Quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	The authors reported that “trace” levels (not quantified) of all OPE diester metabolites were found in procedural blanks. The trace concentrations measured in blanks were subtracted from sample values for each batch. Further details were not provided. Despite this evidence of some contamination, there was nothing to suggest important error, or that the method used to address this issue was biased.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z. Z., Kannan, K., Robinson, M., Foley, H. B., Grubbs, B., Al-Marayati, L., Lerner, D., Lurvey, N., Aung, M. T., Habre, R., Dunton, G. F., Farzan, S. F., Breton, C. V., Bastain, T. M. (2023). Sex-specific effects of prenatal organophosphate ester (OPE) metabolite mixtures and adverse infant birth outcomes in the maternal and developmental risks from environmental and social stressors (MADRES) pregnancy cohort. Environmental Research 226:115703.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, gestational-age adjusted birth weight z-scores
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577122

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	BCEP was analyzed using HPLC with mass spectrometry. Internal standards were used for each OPE. An overview of laboratory methods was provided in limited detail. The authors stated that each batch analyzed included spiked quality control samples and replicates of quality control standard reference materials. Repeated analysis of pooled urine QC samples across batches showed coefficient of variation of ±12–31%. Though precision was variable, there was no evidence of important error.
	Metric 22:	Matrix Adjustment	Medium	Dilution was addressed using specific gravity adjusted OPE exposure variables. There was no rationale provided, but there was no evidence the approach was inappropriate. Results were only presented as matrix-adjusted.

Additional Comments: This study analyzed the relationship between several OPEs and birth outcomes among 421 mother-infant pairs from the Los Angeles MADRES cohort of predominately low-income Latino subjects. BCEP, a major metabolite of TCEP, was measured in a third trimester spot urine sample. The study found no association between urinary BCEP and either gestational age at birth weight for gestational age overall or stratified by infant sex. Strengths of this study include the prospective design. A potential limitation includes variability in gestational age estimated for the same participant based on alternative methods (e.g. ultrasound from different trimesters), but there was no evidence of important error or bias.

Overall Quality Determination **Medium**

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z., Kannan, K., Robinson, M., Foley, H. B., Yang, T., Vigil, M. J., Chen, X., Grubbs, B., Lerner, D., Lurvey, N., Al-Marayati, L., Habre, R., Dunton, G. F., Farzan, S. F., Aung, M. T., Breton, C. V., Bastain, T. M. (2023). Prenatal exposures to organophosphate ester metabolite mixtures and children's neurobehavioral outcomes in the MADRES pregnancy cohort. Environmental Health 22(1):66.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Internalizing Problems, Externalizing Problems, and Total Problems scores from the Child Behavior Checklist (CBCL)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581666

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	This study used data from the ongoing MADRES (Maternal and Developmental Risks from Environmental and Social Stressors) cohort to analyze associations between the TCEP metabolite bis(2-chloroethyl) phosphate (BCEP) in maternal urine and child behavioral outcomes at age 36 months. Urine samples were collected in the third trimester between 2017 and 2019. MADRES includes predominately low-income Hispanic/Latino mother-child pairs recruited through community health practices and advertising in urban Los Angeles. Eligible women are > aged 18 years, have singleton pregnancies, and are recruited prior to 30 weeks' gestation. Exclusion criteria included disabilities, current incarceration, and HIV positivity. Participation rates were not described. However, there was no evidence to suggest that participation was associated with either TCEP exposure or child behavior. A total of n=204 mother-infant pairs were included in the final analysis, compared to the n=426 full cohort of MADRES participants. The study reports that the distribution of OPE metabolites was similar between the two subsets.
	Metric 2:	Attrition	Medium	Of 774 MADRES cohort births through August 2022, 426 had prenatal urinary organophosphate ester (OPE) measures. From this OPE subgroup, the final analysis sample of 204 children excluded 181 children who either had not reached 36 months or had passed that age before the Child Behavior Checklist (CBCL) was added to the study, and another 41 who were missing CBCL data. Descriptive data showed that the analysis sample and urinary OPE subsample had characteristics similar to the parent cohort (e.g. 54 to 56% < high school education, mean gestational age 39 weeks in all groups). TCEP concentrations were also similar in the final analysis sample and OPE subsample (68.6 vs 68.3% detection rates; median [IQR] 0.47 (0.02-1.60) vs. 0.53 (0.02-1.62). Though attrition was high (the analysis sample included 26.3% of the parent cohort), there was no evidence of bias.
	Metric 3:	Comparison Group	Medium	Analyses compared child behavioral outcomes among participants with lower vs. higher BCEP in maternal urine during gestation. Though descriptive data on participant characteristics were not stratified by urinary BCEP concentrations, there was no evidence to suggest the comparison group with lower BCEP differed from those with higher concentrations and key demographic characteristics were considered as potential covariates in statistical analysis.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z., Kannan, K., Robinson, M., Foley, H. B., Yang, T., Vigil, M. J., Chen, X., Grubbs, B., Lerner, D., Lurvey, N., Al-Marayati, L., Habre, R., Dunton, G. F., Farzan, S. F., Aung, M. T., Breton, C. V., Bastain, T. M. (2023). Prenatal exposures to organophosphate ester metabolite mixtures and children's neurobehavioral outcomes in the MADRES pregnancy cohort. Environmental Health 22(1):66.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Internalizing Problems, Externalizing Problems, and Total Problems scores from the Child Behavior Checklist (CBCL)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581666

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	Prenatal TCEP exposure was estimated by BCEP concentrations in a single third trimester urine sample adjusted for dilution using specific gravity. BCEP was detected in 68.6% of samples. In several models, exposure was categorized in tertiles, and values below LOD were used as the referent. Analyses also imputed values below the LOD of 0.02 ng/mL as LOD divided by the square root of 2. Use of the main TCEP metabolite as a biomarker was a strength of exposure measurement. Using a single spot urine sample was a possible limitation: as a single measure may misclassify habitual exposure given uncertain variability in and sources of urinary BCEP in this population. Another potential limitation is that unmetabolized TCEP—which was detected along with BCEP in 37% of subjects in another study – was not measured (e.g., Hou et al. 2020 HEROID 10143372). Use of organophosphate ester measures from late vs early pregnancy is also a possible limitation; other studies found adverse associations with toddler neurobehavioral outcomes only for first trimester measures (e.g., Wang et al 2023 PMID37856202). However, there was no evidence that third trimester concentrations would systematically differ from those in early gestation. While some non-differential misclassification of exposure is likely, there was no evidence of bias.
	Metric 5:	Exposure Levels	Medium	BCEP was analyzed using both natural log-transformed continuous measures and using tertiles. While the range of exposures was relatively small, the range is likely large enough to allow for sufficient contrast (25th-75th percentiles: 0.02-1.60 ng/mL). No concerns.
	Metric 6:	Temporality	Medium	BCEP concentrations were measured prenatally, preceding childbirth and all observations of child behavior. While it is unclear whether the exact relevant time window for developmental effects due to TCEP exposure was captured, there were no specific concerns for temporality.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Child behavioral outcomes were characterized based on maternal responses to the Child Behavior Checklist (CBCL), which was orally administered at 36 months. The CBCL is a validated instrument widely used for evaluating emotional and behavioral problems in children aged 1.5 to 5 years. To facilitate comparisons with previous studies, primary analyses used raw composite scores for three common CBCL metrics: internalizing, externalizing and total problems, adjusting for prematurity-corrected age at CBCL administration and child sex. CBCL T-scores, standardized vs. a US population, were also analyzed. Additional CBCL subscales (e.g. emotionally reactive, attention problems) were not analyzed in this paper. Interviewer training was not discussed, nor was the CBCL validity or reliability within the study population. However, there was no evidence of inadequacies in outcome assessment.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z., Kannan, K., Robinson, M., Foley, H. B., Yang, T., Vigil, M. J., Chen, X., Grubbs, B., Lerner, D., Lurvey, N., Al-Marayati, L., Habre, R., Dunton, G. F., Farzan, S. F., Aung, M. T., Breton, C. V., Bastain, T. M. (2023). Prenatal exposures to organophosphate ester metabolite mixtures and children's neurobehavioral outcomes in the MADRES pregnancy cohort. Environmental Health 22(1):66.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Internalizing Problems, Externalizing Problems, and Total Problems scores from the Child Behavior Checklist (CBCL)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581666

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	High	Results were presented or described for all analyses included as aims.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Covariates were selected based on prior literature and a minimally sufficient set developed using a directed acyclic graph (DAG). Models adjusted for maternal age, parity, pre-pregnancy BMI, race/ethnicity, income, education, child sex, child adjusted age at CBCL administration, and study design or sample collection variables that changed effect estimates by 10% or more (recruitment site, specimen collection season, gestational age at sample collection). Maternal smoking during pregnancy was analyzed by exclusion in a sensitivity analysis due to low prevalence of this behavior. Potential confounding or modification by factors such as size at birth, gestational diabetes or hypertension was not discussed; however, these variables might also be intermediates and thus appropriately excluded. There was no evidence of important bias due to residual confounding.
	Metric 10:	Covariate Characterization	Medium	Covariates were characterized using information from interviewer-administered questionnaires and medical records. Validation was not discussed, but there was no evidence of important error or bias.
	Metric 11:	Co-exposure Counfounding	Medium	Correlations with BCEP were weak (Spearman's r <0.25) for other organophosphate ester (OPE) metabolites measured in this study, which included diphenyl phosphate (DPHP), a composite of di-n-butyl phosphate and di-isobutyl phosphate which co-eluted (DNBP + DIBP), bis(1,3,-dichloro-2-propyl) phosphate (BDCIPP), bis(butoxethyl) phosphate (BBOEP), bis(1-chloro-2-propyl) phosphate (BCIPP), bis(2-ethylhexyl) phosphate (BEHP), bis(2-methylphenyl) phosphate (BMPP), and dipropyl phosphate (DPRP). Co-exposure to other OPEs during pregnancy was evaluated using Bayesian methods to analyze chemical mixtures as well as synergistic interactions among chemicals. Confounding by postnatal exposure to these chemicals was not evaluated. However there was no evidence of bias.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	A prospective cohort design was used to analyze the relationship between prenatal TCEP exposure and child behavior outcomes at age 36 months. The study design was appropriate.
	Metric 13:	Statistical Power	Medium	The sample included 204 mother-child pairs. There was variability in both exposure (BCEP median [IQR] = 0.47 [0.02-1.60] and CBCL outcome variables (e.g., for internalizing problems median 8.0 [IQR = 12.0]). There was no evidence of inadequate statistical power.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what was done and to replicate findings.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z., Kannan, K., Robinson, M., Foley, H. B., Yang, T., Vigil, M. J., Chen, X., Grubbs, B., Lerner, D., Lurvey, N., Al-Marayati, L., Habre, R., Dunton, G. F., Farzan, S. F., Aung, M. T., Breton, C. V., Bastain, T. M. (2023). Prenatal exposures to organophosphate ester metabolite mixtures and children's neurobehavioral outcomes in the MADRES pregnancy cohort. Environmental Health 22(1):66.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Internalizing Problems, Externalizing Problems, and Total Problems scores from the Child Behavior Checklist (CBCL)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581666

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were presented for the sample as a whole. Urinary BCEP was described using percentiles, the geometric means, and the minimum and maximum. To satisfy model assumptions and evaluate potentially non-linear dose-response, primary models for individual exposures analyzed associations with natural-log transformed CBCL scores using either BCEP tertiles or using generalized additive models. Both crude and adjusted effect estimates were presented with 95% confidence intervals. Interaction terms and stratified models were used to assess effect modification by child sex. In addition, Bayesian kernel machine regression (BKMR) models were used to analyze OPE mixtures, and Bayesian semiparametric regression models were used to analyze synergistic or antagonistic interactions among pairs of chemicals. Both sets of models allowed for non-linear dose-response patterns. Results were presented graphically and using posterior inclusion probabilities with 95% credible intervals. Sensitivity analyses to assess robustness included excluding children whose mothers smoked during pregnancy, analyzing CBCL t-scores vs raw scores, and adjusting BKMR smoothness parameters. Analyses appeared to be appropriate and there was no evidence of deficiencies.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP, the primary metabolite of TCEP, were used to estimate exposure.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	BCEP was detected in 68.63% of urine samples; values below the LOD of 0.02 ng/mL were imputed as LOD divided by the square root of two. Though a substantial proportion of samples were below LOD, there was no evidence of inadequate sensitivity.
	Metric 19:	Biomarker Stability	Medium	Processed samples were stored at -80 degrees C prior to shipment. The duration of processing time at room temperature, and total storage time prior to analysis, were not described, but there was no evidence of instability or degradation. Quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	Urine samples were collected in sterile specimen containers, and aliquoted in glass tubes. The authors reported that "trace" levels (not quantified) of all OPE diester metabolites were found in procedural blanks, and that the concentrations measured in blanks were subtracted from sample values for each batch. Further details, such as variability in trace amounts across batches and the number of batches analyzed, were not provided. However, there was no evidence of inappropriate handling of samples, or of bias resulting from the approach used to address trace levels in blanks.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Hernandez-Castro, I., Eckel, S. P., Howe, C. G., Niu, Z., Kannan, K., Robinson, M., Foley, H. B., Yang, T., Vigil, M. J., Chen, X., Grubbs, B., Lerner, D., Lurvey, N., Al-Marayati, L., Habre, R., Dunton, G. F., Farzan, S. F., Aung, M. T., Breton, C. V., Bastain, T. M. (2023). Prenatal exposures to organophosphate ester metabolite mixtures and children's neurobehavioral outcomes in the MADRES pregnancy cohort. Environmental Health 22(1):66.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Internalizing Problems, Externalizing Problems, and Total Problems scores from the Child Behavior Checklist (CBCL)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581666

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	Medium	BCEP was measured using HPLC with mass spectrometry. Unspecified internal standards were used. Each batch analyzed included spiked quality control samples and replicates of quality control standard reference materials. Standardized reference materials had coefficients of variation of up to ± 12–40%; average recoveries of matrix spiked samples ranged from 70.4–133% (coefficients of variation ± 9–19%). Though details were limited and precision was variable, there was no evidence of important error or bias in BCEP measures.
	Metric 22:	Matrix Adjustment	Medium	Samples were adjusted for dilution using specific gravity measured in room temperature urine samples using a digital handheld refractometer. Specific gravity-adjusted concentrations were used in analysis. A rationale for this approach was not discussed, but there was no evidence for concern. Results were only shown for concentrations adjusted for dilution.

Additional Comments: This study used data on 204 mother-infant pairs from the Los Angeles MADRES cohort of predominately low-income Latino subjects. TCEP exposure was estimated using BCEP in a single third trimester spot urine sample collected in 2017-2019; child behavior was assessed using scores on the Child Behavior Checklist at age 36 months. The study found positive, non-significant associations between BCEP and CBCL internalizing, externalizing and total problems scores. Sex differences were not significant. In addition, using BKMR models, there were stronger positive associations between DNBP+DIBP (co-eluting OPEs) and both internalizing and total problems scores at higher quartiles of BCEP. A significant interaction, tested for the strongest interaction, was confirmed using generalized additive models. Strengths of this study include the prospective design, and that characteristics of the analysis sample were similar to the parent cohort.Potential limitations include that the analysis sample comprised 204 of more than 700 MADRES participants, and that exposure estimation was based on a single spot sample from the third trimester. There was no evidence of important error or bias.

Overall Quality Determination **Medium**

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants were drawn from the National Health and Nutrition Examination Survey (NHANES) study (2011-2012 cycle), a study on a nationally representative sample conducted by the CDC. This cycle included both urinary measures of organophosphate flame retardants and lung function spirometry measures. All NHANES participants aged 20 years or older (n=5,560) were eligible if they had data on urinary organophosphate flame retardant (OPFR) metabolites and pulmonary parameters. From the 5,560 adults included in the 2011-2012 cycle, 4,178 without OPFR metabolites were excluded, 384 with missing or inadequate pulmonary parameters were excluded, and 11 pregnant women were excluded. This resulted in a final sample of 987 participants. No information was provided to demonstrate the exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study, however, the study is intended to be a nationally representative sample and risk of selection bias is low. There was nothing to suggest the subset with available OPFR measures was selective. Further details on recruitment and participant selection can be found on the CDC NHANES website.
	Metric 2:	Attrition	Medium	Reasons for exclusion were appropriately outlined, and only those with complete exposure and outcome information were included in the analysis. Details such as exclusions due to ineligibility (e.g., due to heart attack or stroke) vs. inadequate quality of spirometry were not provided, but there was no evidence of bias.
	Metric 3:	Comparison Group	High	All participants were drawn from NHANES, a nationally representative study sample, using the same inclusion and exclusion criteria. The study was restricted to adults (>=20 years old) and pregnant women were excluded.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	A metabolite of TCEP, BCEP, was measured in urine to indicate exposure to TCEP. Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry. Limits of detection and QA/QC procedures were described, and further details can be found on the CDC NHANES website. The extent to which chronic TCEP exposure may be misclassified using a single spot urine concentration of BCEP is uncertain. The authors did not discuss the half-life of urinary BCEP or include measures of the parent compound along with the metabolite (Wang et al 2020 HEROID 7276658; Hou et al 10143372).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	Exposure to TCEP was evaluated based on urinary BCEP concentrations, analyzed continuously. BCEP was detectable in 85.31% of urine samples, and values below LOD were imputed as the LOD divided by the square root of 2. No concerns about exposure distribution (BCEP median [25th-75th percentile]: 0.51 ug/L [0.21, 1.03]). Urinary creatinine was included in models to account for urine dilution.
	Metric 6:	Temporality	Medium	Urine samples used to estimate OPFR exposure and spirometry used to characterize the respiratory function outcomes were collected during the same NHANES study visit. While temporality cannot be established due to the cross-sectional design given the short half-life (<30 days) of the exposure biomarker, there is no evidence of reverse causation.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Spirometry testing procedures were performed according to the recommendations of the American Thoracic Society (ATS). Measurements included forced expiratory volume first second (FEV1), forced vital capacity (FVC), peak expiratory flowrate (PEF), forced expiratory flow at 25−75% of the FVC (FEF25−75%), and FEV1/FVC. Measurements were performed in a standing position (except for physical limitation). "A maximum of eight spirometry tests were conducted to obtain three acceptable curves. The two highest values for FEV1 and FVC were used for quality ratings." Only those subjects with FEV1 and FVC measurements rated as A or B (according to ATS data collection standards) were included. Further details can be found in the NHANES Respiratory Health Spirometry Procedures Manual. Blinding was not mentioned, however, spirometry data was collected at the same visit as urine was collected, indirectly indicating spirometry technicians would be unaware of exposure status.
	Metric 8:	Reporting Bias	High	All outcomes outlined in the abstract, introduction, and methods were provided in the results or supplementary materials. Risk estimates were reported with p-values and confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study authors provided a simple model adjusted for age, sex, and race (model 1) and a model additionally adjusted for BMI, serum cotinine, smoking status, physical activity, family poverty/income ratio, educational level, and urinary creatinine. Study authors conducted a stratified analysis to evaluate interaction by smoking status (never, former, current). Examining the robustness of findings in sensitivity analyses was not discussed (e.g. potential influence of participants with diagnoses respiratory illness such as asthma).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were collected by trained interviewers during NHANES study visits; these included self-reported sociodemographic measures. Complete information on collection covariates can be found in the NHANES (2011-2012) Procedure Manuals. There was no evidence of covariates with inadequate validity.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPFR metabolites were measured in urine. However, correlations among metabolites were not provided, and models did not adjust for other metabolites. There was no direct evidence of confounding by these or other co-exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was appropriate to examine the research question (i.e., the association between exposure to TCEP and pulmonary function). Appropriate statistical methods (linear regression models) were used to evaluate the exposure-outcome relationship.
	Metric 13:	Statistical Power	Medium	There was variability in exposure, and the number of participants (n=987) was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	High	Linear regression models were adequately described. Study authors noted that exposure data was natural log transformed to normalize the distribution. Minor potential concerns include that the non-linearity of dose-response relationships and sex-specific effects were not discussed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Supplemental material (Table S1) and the Pubchem dashboard indicate BCEP is the direct metabolite of TCEP. No concerns about accuracy or precision of metabolite measurements in urine. However, the extent to which a single spot urine BCEP accurately classified TCEP exposure in non-occupationally exposed populations is not fully established.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD (0.10 ug/L) was low enough to detect chemicals in a sufficient percentage of the samples (85.31%) to address the research question, and the method was adequately described.
	Metric 19:	Biomarker Stability	Medium	Samples were collected during an NHANES study visit and documentation on storage is provided in the NHANES Laboratory Procedures Manual. There was no information provided on biomarker stability, but no evidence for concern was documented.
	Metric 20:	Sample Contamination	Medium	There is no specific information on sample contamination, but study authors note that samples were collected ”according to the analytic guidelines of the NHANES.” Specific information may be found in the NHANES Laboratory Procedures Manual.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Model 1 adjusted for age, sex and race but excluded urinary creatinine, while Model 2 provides additionally adjusted for urinary creatinine and several other covariates. Associations with and without creatinine were shown. However, the influence of creatinine adjustment vs other confounding cannot be isolated.

Additional Comments: This was a well-conducted study based on available NHANES data in adults, using the 2011-2012 cycle. There were no large concerns with participant selection or methods relating to exposure measurement, outcome ascertainment, and statistical analysis. Minor concerns include that temporality cannot be fully established due to the cross-sectional design, and that co-exposure confounding was not explored.

Overall Quality Determination **Medium**

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants were drawn from the National Health and Nutrition Examination Survey (NHANES) study (2011-2012 cycle), a study on a nationally representative sample conducted by the CDC. This cycle included both urinary measures of organophosphate flame retardants and lung function spirometry measures. All NHANES participants aged 20 years or older (n=5,560) were eligible if they had data on urinary organophosphate flame retardant (OPFR) metabolites and pulmonary parameters. From the 5,560 adults included in the 2011-2012 cycle, 4,178 without OPFR metabolites were excluded, 384 with missing or inadequate pulmonary parameters were excluded, and 11 pregnant women were excluded. This resulted in a final sample of 987 participants. No information was provided to demonstrate the exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study, however, the study is intended to be a nationally representative sample and risk of selection bias is low. There was nothing to suggest the subset with available OPFR measures was selective. Further details on recruitment and participant selection can be found on the CDC NHANES website.
	Metric 2:	Attrition	Medium	Reasons for exclusion were appropriately outlined, and only those with complete exposure and outcome information were included in the analysis. Details such as exclusions due to ineligibility (e.g., due to heart attack or stroke) vs. inadequate quality of spirometry were not provided, but there was no evidence of bias.
	Metric 3:	Comparison Group	High	All participants were drawn from NHANES, a nationally representative study sample, using the same inclusion and exclusion criteria. The study was restricted to adults (>=20 years old) and pregnant women were excluded.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	A metabolite of TCEP, BCEP, was measured in urine to indicate exposure to TCEP. Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry. Limits of detection and QA/QC procedures were described, and further details can be found on the CDC NHANES website. The extent to which chronic TCEP exposure may be misclassified using a single spot urine concentration of BCEP is uncertain. The authors did not discuss the half-life of urinary BCEP or include measures of the parent compound along with the metabolite (Wang et al 2020 HEROID 7276658; Hou et al 10143372).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	Exposure to TCEP was evaluated based on urinary BCEP concentrations, analyzed continuously. BCEP was detectable in 85.31% of urine samples, and values below LOD were imputed as the LOD divided by the square root of 2. No concerns about exposure distribution (BCEP median [25th-75th percentile]: 0.51 ug/L [0.21, 1.03]). Urinary creatinine was included in models to account for urine dilution.
	Metric 6:	Temporality	Medium	Urine samples used to estimate OPFR exposure and spirometry used to characterize the respiratory function outcomes were collected during the same NHANES study visit. While temporality cannot be established due to the cross-sectional design given the short half-life (<30 days) of the exposure biomarker, there is no evidence of reverse causation.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Spirometry testing procedures were performed according to the recommendations of the American Thoracic Society (ATS). Measurements included forced expiratory volume first second (FEV1), forced vital capacity (FVC), peak expiratory flowrate (PEF), forced expiratory flow at 25−75% of the FVC (FEF25−75%), and FEV1/FVC. Measurements were performed in a standing position (except for physical limitation). "A maximum of eight spirometry tests were conducted to obtain three acceptable curves. The two highest values for FEV1 and FVC were used for quality ratings." Only those subjects with FEV1 and FVC measurements rated as A or B (according to ATS data collection standards) were included. Further details can be found in the NHANES Respiratory Health Spirometry Procedures Manual. Blinding was not mentioned, however, spirometry data was collected at the same visit as urine was collected, indirectly indicating spirometry technicians would be unaware of exposure status.
	Metric 8:	Reporting Bias	High	All outcomes outlined in the abstract, introduction, and methods were provided in the results or supplementary materials. Risk estimates were reported with p-values and confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study authors provided a simple model adjusted for age, sex, and race (model 1) and a model additionally adjusted for BMI, serum cotinine, smoking status, physical activity, family poverty/income ratio, educational level, and urinary creatinine. Study authors conducted a stratified analysis to evaluate interaction by smoking status (never, former, current). Examining the robustness of findings in sensitivity analyses was not discussed (e.g. potential influence of participants with diagnoses respiratory illness such as asthma).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were collected by trained interviewers during NHANES study visits; these included self-reported sociodemographic measures. Complete information on collection covariates can be found in the NHANES (2011-2012) Procedure Manuals. There was no evidence of covariates with inadequate validity.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPFR metabolites were measured in urine. However, correlations among metabolites were not provided, and models did not adjust for other metabolites. There was no direct evidence of confounding by these or other co-exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was appropriate to examine the research question (i.e., the association between exposure to TCEP and pulmonary function). Appropriate statistical methods (linear regression models) were used to evaluate the exposure-outcome relationship.
	Metric 13:	Statistical Power	Medium	There was variability in exposure, and the number of participants (n=987) was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	High	Linear regression models were adequately described. Study authors noted that exposure data was natural log transformed to normalize the distribution. Minor potential concerns include that the non-linearity of dose-response relationships and sex-specific effects were not discussed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Supplemental material (Table S1) and the Pubchem dashboard indicate BCEP is the direct metabolite of TCEP. No concerns about accuracy or precision of metabolite measurements in urine. However, the extent to which a single spot urine BCEP accurately classified TCEP exposure in non-occupationally exposed populations is not fully established.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD (0.10 ug/L) was low enough to detect chemicals in a sufficient percentage of the samples (85.31%) to address the research question, and the method was adequately described.
	Metric 19:	Biomarker Stability	Medium	Samples were collected during an NHANES study visit and documentation on storage is provided in the NHANES Laboratory Procedures Manual. There was no information provided on biomarker stability, but no evidence for concern was documented.
	Metric 20:	Sample Contamination	Medium	There is no specific information on sample contamination, but study authors note that samples were collected ”according to the analytic guidelines of the NHANES.” Specific information may be found in the NHANES Laboratory Procedures Manual.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Model 1 adjusted for age, sex and race but excluded urinary creatinine, while Model 2 provides additionally adjusted for urinary creatinine and several other covariates. Associations with and without creatinine were shown. However, the influence of creatinine adjustment vs other confounding cannot be isolated.

Additional Comments: This was a well-conducted study based on available NHANES data in adults, using the 2011-2012 cycle. There were no large concerns with participant selection or methods relating to exposure measurement, outcome ascertainment, and statistical analysis. Minor concerns include that temporality cannot be fully established due to the cross-sectional design, and that co-exposure confounding was not explored.

Overall Quality Determination **Medium**

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants were drawn from the National Health and Nutrition Examination Survey (NHANES) study (2011-2012 cycle), a study on a nationally representative sample conducted by the CDC. This cycle included both urinary measures of organophosphate flame retardants and lung function spirometry measures. All NHANES participants aged 20 years or older (n=5,560) were eligible if they had data on urinary organophosphate flame retardant (OPFR) metabolites and pulmonary parameters. From the 5,560 adults included in the 2011-2012 cycle, 4,178 without OPFR metabolites were excluded, 384 with missing or inadequate pulmonary parameters were excluded, and 11 pregnant women were excluded. This resulted in a final sample of 987 participants. No information was provided to demonstrate the exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study, however, the study is intended to be a nationally representative sample and risk of selection bias is low. There was nothing to suggest the subset with available OPFR measures was selective. Further details on recruitment and participant selection can be found on the CDC NHANES website.
	Metric 2:	Attrition	Medium	Reasons for exclusion were appropriately outlined, and only those with complete exposure and outcome information were included in the analysis. Details such as exclusions due to ineligibility (e.g., due to heart attack or stroke) vs. inadequate quality of spirometry were not provided, but there was no evidence of bias.
	Metric 3:	Comparison Group	High	All participants were drawn from NHANES, a nationally representative study sample, using the same inclusion and exclusion criteria. The study was restricted to adults (>=20 years old) and pregnant women were excluded.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	A metabolite of TCEP, BCEP, was measured in urine to indicate exposure to TCEP. Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry. Limits of detection and QA/QC procedures were described, and further details can be found on the CDC NHANES website. The extent to which chronic TCEP exposure may be misclassified using a single spot urine concentration of BCEP is uncertain. The authors did not discuss the half-life of urinary BCEP or include measures of the parent compound along with the metabolite (Wang et al 2020 HEROID 7276658; Hou et al 10143372).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	Exposure to TCEP was evaluated based on urinary BCEP concentrations, analyzed continuously. BCEP was detectable in 85.31% of urine samples, and values below LOD were imputed as the LOD divided by the square root of 2. No concerns about exposure distribution (BCEP median [25th-75th percentile]: 0.51 ug/L [0.21, 1.03]). Urinary creatinine was included in models to account for urine dilution.
	Metric 6:	Temporality	Medium	Urine samples used to estimate OPFR exposure and spirometry used to characterize the respiratory function outcomes were collected during the same NHANES study visit. While temporality cannot be established due to the cross-sectional design given the short half-life (<30 days) of the exposure biomarker, there is no evidence of reverse causation.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Spirometry testing procedures were performed according to the recommendations of the American Thoracic Society (ATS). Measurements included forced expiratory volume first second (FEV1), forced vital capacity (FVC), peak expiratory flowrate (PEF), forced expiratory flow at 25−75% of the FVC (FEF25−75%), and FEV1/FVC. Measurements were performed in a standing position (except for physical limitation). "A maximum of eight spirometry tests were conducted to obtain three acceptable curves. The two highest values for FEV1 and FVC were used for quality ratings." Only those subjects with FEV1 and FVC measurements rated as A or B (according to ATS data collection standards) were included. Further details can be found in the NHANES Respiratory Health Spirometry Procedures Manual. Blinding was not mentioned, however, spirometry data was collected at the same visit as urine was collected, indirectly indicating spirometry technicians would be unaware of exposure status.
	Metric 8:	Reporting Bias	High	All outcomes outlined in the abstract, introduction, and methods were provided in the results or supplementary materials. Risk estimates were reported with p-values and confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study authors provided a simple model adjusted for age, sex, and race (model 1) and a model additionally adjusted for BMI, serum cotinine, smoking status, physical activity, family poverty/income ratio, educational level, and urinary creatinine. Study authors conducted a stratified analysis to evaluate interaction by smoking status (never, former, current). Examining the robustness of findings in sensitivity analyses was not discussed (e.g. potential influence of participants with diagnoses respiratory illness such as asthma).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were collected by trained interviewers during NHANES study visits; these included self-reported sociodemographic measures. Complete information on collection covariates can be found in the NHANES (2011-2012) Procedure Manuals. There was no evidence of covariates with inadequate validity.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPFR metabolites were measured in urine. However, correlations among metabolites were not provided, and models did not adjust for other metabolites. There was no direct evidence of confounding by these or other co-exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was appropriate to examine the research question (i.e., the association between exposure to TCEP and pulmonary function). Appropriate statistical methods (linear regression models) were used to evaluate the exposure-outcome relationship.
	Metric 13:	Statistical Power	Medium	There was variability in exposure, and the number of participants (n=987) was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	High	Linear regression models were adequately described. Study authors noted that exposure data was natural log transformed to normalize the distribution. Minor potential concerns include that the non-linearity of dose-response relationships and sex-specific effects were not discussed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Supplemental material (Table S1) and the Pubchem dashboard indicate BCEP is the direct metabolite of TCEP. No concerns about accuracy or precision of metabolite measurements in urine. However, the extent to which a single spot urine BCEP accurately classified TCEP exposure in non-occupationally exposed populations is not fully established.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD (0.10 ug/L) was low enough to detect chemicals in a sufficient percentage of the samples (85.31%) to address the research question, and the method was adequately described.
	Metric 19:	Biomarker Stability	Medium	Samples were collected during an NHANES study visit and documentation on storage is provided in the NHANES Laboratory Procedures Manual. There was no information provided on biomarker stability, but no evidence for concern was documented.
	Metric 20:	Sample Contamination	Medium	There is no specific information on sample contamination, but study authors note that samples were collected ”according to the analytic guidelines of the NHANES.” Specific information may be found in the NHANES Laboratory Procedures Manual.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Model 1 adjusted for age, sex and race but excluded urinary creatinine, while Model 2 provides additionally adjusted for urinary creatinine and several other covariates. Associations with and without creatinine were shown. However, the influence of creatinine adjustment vs other confounding cannot be isolated.

Additional Comments: This was a well-conducted study based on available NHANES data in adults, using the 2011-2012 cycle. There were no large concerns with participant selection or methods relating to exposure measurement, outcome ascertainment, and statistical analysis. Minor concerns include that temporality cannot be fully established due to the cross-sectional design, and that co-exposure confounding was not explored.

Overall Quality Determination **Medium**

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants were drawn from the National Health and Nutrition Examination Survey (NHANES) study (2011-2012 cycle), a study on a nationally representative sample conducted by the CDC. This cycle included both urinary measures of organophosphate flame retardants and lung function spirometry measures. All NHANES participants aged 20 years or older (n=5,560) were eligible if they had data on urinary organophosphate flame retardant (OPFR) metabolites and pulmonary parameters. From the 5,560 adults included in the 2011-2012 cycle, 4,178 without OPFR metabolites were excluded, 384 with missing or inadequate pulmonary parameters were excluded, and 11 pregnant women were excluded. This resulted in a final sample of 987 participants. No information was provided to demonstrate the exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study, however, the study is intended to be a nationally representative sample and risk of selection bias is low. There was nothing to suggest the subset with available OPFR measures was selective. Further details on recruitment and participant selection can be found on the CDC NHANES website.
	Metric 2:	Attrition	Medium	Reasons for exclusion were appropriately outlined, and only those with complete exposure and outcome information were included in the analysis. Details such as exclusions due to ineligibility (e.g., due to heart attack or stroke) vs. inadequate quality of spirometry were not provided, but there was no evidence of bias.
	Metric 3:	Comparison Group	High	All participants were drawn from NHANES, a nationally representative study sample, using the same inclusion and exclusion criteria. The study was restricted to adults (>=20 years old) and pregnant women were excluded.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	A metabolite of TCEP, BCEP, was measured in urine to indicate exposure to TCEP. Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry. Limits of detection and QA/QC procedures were described, and further details can be found on the CDC NHANES website. The extent to which chronic TCEP exposure may be misclassified using a single spot urine concentration of BCEP is uncertain. The authors did not discuss the half-life of urinary BCEP or include measures of the parent compound along with the metabolite (Wang et al 2020 HEROID 7276658; Hou et al 10143372).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	Exposure to TCEP was evaluated based on urinary BCEP concentrations, analyzed continuously. BCEP was detectable in 85.31% of urine samples, and values below LOD were imputed as the LOD divided by the square root of 2. No concerns about exposure distribution (BCEP median [25th-75th percentile]: 0.51 ug/L [0.21, 1.03]). Urinary creatinine was included in models to account for urine dilution.
	Metric 6:	Temporality	Medium	Urine samples used to estimate OPFR exposure and spirometry used to characterize the respiratory function outcomes were collected during the same NHANES study visit. While temporality cannot be established due to the cross-sectional design given the short half-life (<30 days) of the exposure biomarker, there is no evidence of reverse causation.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Spirometry testing procedures were performed according to the recommendations of the American Thoracic Society (ATS). Measurements included forced expiratory volume first second (FEV1), forced vital capacity (FVC), peak expiratory flowrate (PEF), forced expiratory flow at 25−75% of the FVC (FEF25−75%), and FEV1/FVC. Measurements were performed in a standing position (except for physical limitation). "A maximum of eight spirometry tests were conducted to obtain three acceptable curves. The two highest values for FEV1 and FVC were used for quality ratings." Only those subjects with FEV1 and FVC measurements rated as A or B (according to ATS data collection standards) were included. Further details can be found in the NHANES Respiratory Health Spirometry Procedures Manual. Blinding was not mentioned, however, spirometry data was collected at the same visit as urine was collected, indirectly indicating spirometry technicians would be unaware of exposure status.
	Metric 8:	Reporting Bias	High	All outcomes outlined in the abstract, introduction, and methods were provided in the results or supplementary materials. Risk estimates were reported with p-values and confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study authors provided a simple model adjusted for age, sex, and race (model 1) and a model additionally adjusted for BMI, serum cotinine, smoking status, physical activity, family poverty/income ratio, educational level, and urinary creatinine. Study authors conducted a stratified analysis to evaluate interaction by smoking status (never, former, current). Examining the robustness of findings in sensitivity analyses was not discussed (e.g. potential influence of participants with diagnoses respiratory illness such as asthma).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were collected by trained interviewers during NHANES study visits; these included self-reported sociodemographic measures. Complete information on collection covariates can be found in the NHANES (2011-2012) Procedure Manuals. There was no evidence of covariates with inadequate validity.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPFR metabolites were measured in urine. However, correlations among metabolites were not provided, and models did not adjust for other metabolites. There was no direct evidence of confounding by these or other co-exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was appropriate to examine the research question (i.e., the association between exposure to TCEP and pulmonary function). Appropriate statistical methods (linear regression models) were used to evaluate the exposure-outcome relationship.
	Metric 13:	Statistical Power	Medium	There was variability in exposure, and the number of participants (n=987) was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	High	Linear regression models were adequately described. Study authors noted that exposure data was natural log transformed to normalize the distribution. Minor potential concerns include that the non-linearity of dose-response relationships and sex-specific effects were not discussed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Supplemental material (Table S1) and the Pubchem dashboard indicate BCEP is the direct metabolite of TCEP. No concerns about accuracy or precision of metabolite measurements in urine. However, the extent to which a single spot urine BCEP accurately classified TCEP exposure in non-occupationally exposed populations is not fully established.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD (0.10 ug/L) was low enough to detect chemicals in a sufficient percentage of the samples (85.31%) to address the research question, and the method was adequately described.
	Metric 19:	Biomarker Stability	Medium	Samples were collected during an NHANES study visit and documentation on storage is provided in the NHANES Laboratory Procedures Manual. There was no information provided on biomarker stability, but no evidence for concern was documented.
	Metric 20:	Sample Contamination	Medium	There is no specific information on sample contamination, but study authors note that samples were collected ”according to the analytic guidelines of the NHANES.” Specific information may be found in the NHANES Laboratory Procedures Manual.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Model 1 adjusted for age, sex and race but excluded urinary creatinine, while Model 2 provides additionally adjusted for urinary creatinine and several other covariates. Associations with and without creatinine were shown. However, the influence of creatinine adjustment vs other confounding cannot be isolated.

Additional Comments: This was a well-conducted study based on available NHANES data in adults, using the 2011-2012 cycle. There were no large concerns with participant selection or methods relating to exposure measurement, outcome ascertainment, and statistical analysis. Minor concerns include that temporality cannot be fully established due to the cross-sectional design, and that co-exposure confounding was not explored.

Overall Quality Determination **Medium**

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants were drawn from the National Health and Nutrition Examination Survey (NHANES) study (2011-2012 cycle), a study on a nationally representative sample conducted by the CDC. This cycle included both urinary measures of organophosphate flame retardants and lung function spirometry measures. All NHANES participants aged 20 years or older (n=5,560) were eligible if they had data on urinary organophosphate flame retardant (OPFR) metabolites and pulmonary parameters. From the 5,560 adults included in the 2011-2012 cycle, 4,178 without OPFR metabolites were excluded, 384 with missing or inadequate pulmonary parameters were excluded, and 11 pregnant women were excluded. This resulted in a final sample of 987 participants. No information was provided to demonstrate the exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study, however, the study is intended to be a nationally representative sample and risk of selection bias is low. There was nothing to suggest the subset with available OPFR measures was selective. Further details on recruitment and participant selection can be found on the CDC NHANES website.
	Metric 2:	Attrition	Medium	Reasons for exclusion were appropriately outlined, and only those with complete exposure and outcome information were included in the analysis. Details such as exclusions due to ineligibility (e.g., due to heart attack or stroke) vs. inadequate quality of spirometry were not provided, but there was no evidence of bias.
	Metric 3:	Comparison Group	High	All participants were drawn from NHANES, a nationally representative study sample, using the same inclusion and exclusion criteria. The study was restricted to adults (>=20 years old) and pregnant women were excluded.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	A metabolite of TCEP, BCEP, was measured in urine to indicate exposure to TCEP. Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry. Limits of detection and QA/QC procedures were described, and further details can be found on the CDC NHANES website. The extent to which chronic TCEP exposure may be misclassified using a single spot urine concentration of BCEP is uncertain. The authors did not discuss the half-life of urinary BCEP or include measures of the parent compound along with the metabolite (Wang et al 2020 HEROID 7276658; Hou et al 10143372).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	Exposure to TCEP was evaluated based on urinary BCEP concentrations, analyzed continuously. BCEP was detectable in 85.31% of urine samples, and values below LOD were imputed as the LOD divided by the square root of 2. No concerns about exposure distribution (BCEP median [25th-75th percentile]: 0.51 ug/L [0.21, 1.03]). Urinary creatinine was included in models to account for urine dilution.
	Metric 6:	Temporality	Medium	Urine samples used to estimate OPFR exposure and spirometry used to characterize the respiratory function outcomes were collected during the same NHANES study visit. While temporality cannot be established due to the cross-sectional design given the short half-life (<30 days) of the exposure biomarker, there is no evidence of reverse causation.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Spirometry testing procedures were performed according to the recommendations of the American Thoracic Society (ATS). Measurements included forced expiratory volume first second (FEV1), forced vital capacity (FVC), peak expiratory flowrate (PEF), forced expiratory flow at 25−75% of the FVC (FEF25−75%), and FEV1/FVC. Measurements were performed in a standing position (except for physical limitation). "A maximum of eight spirometry tests were conducted to obtain three acceptable curves. The two highest values for FEV1 and FVC were used for quality ratings." Only those subjects with FEV1 and FVC measurements rated as A or B (according to ATS data collection standards) were included. Further details can be found in the NHANES Respiratory Health Spirometry Procedures Manual. Blinding was not mentioned, however, spirometry data was collected at the same visit as urine was collected, indirectly indicating spirometry technicians would be unaware of exposure status.
	Metric 8:	Reporting Bias	High	All outcomes outlined in the abstract, introduction, and methods were provided in the results or supplementary materials. Risk estimates were reported with p-values and confidence intervals.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Study authors provided a simple model adjusted for age, sex, and race (model 1) and a model additionally adjusted for BMI, serum cotinine, smoking status, physical activity, family poverty/income ratio, educational level, and urinary creatinine. Study authors conducted a stratified analysis to evaluate interaction by smoking status (never, former, current). Examining the robustness of findings in sensitivity analyses was not discussed (e.g. potential influence of participants with diagnoses respiratory illness such as asthma).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were collected by trained interviewers during NHANES study visits; these included self-reported sociodemographic measures. Complete information on collection covariates can be found in the NHANES (2011-2012) Procedure Manuals. There was no evidence of covariates with inadequate validity.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPFR metabolites were measured in urine. However, correlations among metabolites were not provided, and models did not adjust for other metabolites. There was no direct evidence of confounding by these or other co-exposures.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was appropriate to examine the research question (i.e., the association between exposure to TCEP and pulmonary function). Appropriate statistical methods (linear regression models) were used to evaluate the exposure-outcome relationship.
	Metric 13:	Statistical Power	Medium	There was variability in exposure, and the number of participants (n=987) was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	High	Linear regression models were adequately described. Study authors noted that exposure data was natural log transformed to normalize the distribution. Minor potential concerns include that the non-linearity of dose-response relationships and sex-specific effects were not discussed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Supplemental material (Table S1) and the Pubchem dashboard indicate BCEP is the direct metabolite of TCEP. No concerns about accuracy or precision of metabolite measurements in urine. However, the extent to which a single spot urine BCEP accurately classified TCEP exposure in non-occupationally exposed populations is not fully established.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The LOD (0.10 ug/L) was low enough to detect chemicals in a sufficient percentage of the samples (85.31%) to address the research question, and the method was adequately described.
	Metric 19:	Biomarker Stability	Medium	Samples were collected during an NHANES study visit and documentation on storage is provided in the NHANES Laboratory Procedures Manual. There was no information provided on biomarker stability, but no evidence for concern was documented.
	Metric 20:	Sample Contamination	Medium	There is no specific information on sample contamination, but study authors note that samples were collected ”according to the analytic guidelines of the NHANES.” Specific information may be found in the NHANES Laboratory Procedures Manual.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Zhu, H., Zhang, H., Lu, K., Yang, S., Tang, X., Zhou, M., Sun, G., Zhang, Z., Chu, H. (2022). Chlorinated Organophosphate Flame Retardants Impair the Lung Function via the IL-6/JAK/STAT Signaling Pathway. Environmental Science & Technology 56(24):17858-17869.
Health Outcome(s) and Reported Health Effect(s):	Lung/Respiratory-Spirometry measures: Forced expiratory volume first second (FEV1), Forced vital capacity (FVC), Peak expiratory flow (PEF), Forced expiratory flow at 25-75% of FVC (FEF25-75%), FEV1/FVC
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11365039

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	Samples were analyzed using automated off-line solid phase extraction, reversed-phase high-performance liquid chromatography followed by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Model 1 adjusted for age, sex and race but excluded urinary creatinine, while Model 2 provides additionally adjusted for urinary creatinine and several other covariates. Associations with and without creatinine were shown. However, the influence of creatinine adjustment vs other confounding cannot be isolated.

Additional Comments: This was a well-conducted study based on available NHANES data in adults, using the 2011-2012 cycle. There were no large concerns with participant selection or methods relating to exposure measurement, outcome ascertainment, and statistical analysis. Minor concerns include that temporality cannot be fully established due to the cross-sectional design, and that co-exposure confounding was not explored.

Overall Quality Determination **Medium**

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	This prospective birth cohort study used data from the Health Outcomes and Measures of the Environment (HOME) study, which was conducted in the Greater Cincinnati Metropolitan Area, Ohio. Pregnant women were recruited from March 2003-January 2006 Participants were recruited "using the medical scheduling systems of nine prenatal practices affiliated with three hospitals," with eligibility determined "using clinic records and phone interviews with women" (Braun et al., 2017; HEROID: 6749104). Inclusion criteria were: 1) age >18 years, 2) 13–19 weeks pregnancy, 3) residing in a home built in or before 1978, 5) fluent in English, 6) planning to live in the study area for the next year, 7) planning to continue prenatal care and deliver at the participating hospitals. Exclusion criteria were: living in a mobile or trailer home, were on medications for thyroid or seizure disorders, or diagnosed with bipolar disorder, schizophrenia, diabetes, or cancer requiring radiation or chemotherapy. The participation rate at each step was reported in Braun et al. 2017. From a sampling frame of 8878 women, 1263 were eligible and 340 women were included in the final analytic sample. The study compares those included in the original cohort (n=389) and those finally included in the present analysis (n=340) and reports that maternal characteristics were comparable between the two groups. There is no evidence to suggest substantial selection bias and key elements of the study design were reported.
	Metric 2:	Attrition	Medium	Details on attrition are available in Braun et al. 2017 (HEROID: 6749104). Of 468 women enrolled, 67 dropped out before delivery due to concerns about the time commitment, having family members unwilling to participate, or loss of contact. 389 women delivered singleton births, but only 340 were included in the final analysis for having no congenital malformations, having one spot urine sample quantified for OPE metabolites, and having neonatal anthropometry abstracted from medical records. While it is not clear why these measures were not available for all 389 women, exclusion from the analysis sample is appropriate and there is no evidence that any attrition in the study was related to exposure and outcome.
	Metric 3:	Comparison Group	High	The study reports on demographic characteristics and reports whether differences in characteristics were statistically significant by exposure and outcome variables. While significant differences were reported, these factors were considered as potential covariates in statistical analysis. There is no evidence to suggest that participants were selected differently by outcome or exposure group.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was assessed using an established biomarker of TCEP exposure (the urinary metabolite BCEP) in urine samples collected from mothers twice during pregnancy at an average of 16 $\pm$ 2 and 26 $\pm$ 3 weeks of gestation. BCEP was quantified using high-performance liquid chromatography with tandem mass spectrometry. LOD information is reported. Overall, exposure to TCEP was well-conducted and the use of two measurements increases the certainty in exposure concentrations across pregnancy.
	Metric 5:	Exposure Levels	Medium	The range of exposure is adequate, with no evidence of insufficient variability to estimate associations. BCEP was characterized into tertiles and also analyzed as continuous variable.
	Metric 6:	Temporality	Medium	The study includes measurements of urinary metabolites at two time points: 16 and 26 weeks, for pregnancy outcomes (gestational age and neonatal anthropometric measures). Temporality is therefore clearly established, with exposure preceding outcome. While the exposure is measured during pregnancy both time points captured are during the second trimester, meaning the first trimester, a relevant time period and critical window for fetal development was missed. Additionally, it is unclear whether the second trimester is the most etiologically relevant time window for developmental effects of BCEP.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Medical charts were used to obtain gestational age and infant anthropometric parameters at birth (birth weight (g), length (cm), and head circumference (cm). Standardized birth weight, length, and head circumference z-scores were calculated using values from the 2010 Olsen growth charts. Ponderal index (PI) was calculated as 100x (weight/length). One limitation is that gestational age was estimated by last menstrual period (LMP) for 330 participants, but by other measures (ultrasound (n = 7) or Ballard scores (n = 2) for very limited cases). While it is a widely used method to estimate gestational age, LMP has limitations as it relies on participant recall and also assumes regular 28-day menstrual cycles with conception occurring on the 14th day.Preterm birth was defined as birth prior to 37 completed weeks of gestation.
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed and reported as described in the study aims. Effect estimates are reported with confidence intervals. While the number of exposed/unexposed are not reported in each analysis, the number of observations used are reported consistently for each analysis.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were identified a priori through use of a directed acyclic graph (DAG). Considered covariates included household income, infant sex, marital status, maternal age, maternal blood lead, maternal serum cotinine, maternal education, maternal race, parity, pre-pregnancy BMI. Effect modification by sex was also explored. Key covariates were thus considered, and potential residual confounding is unlikely to have a significant impact on effect estimates.
	Metric 10:	Covariate Characterization	Medium	The exact methodology for obtaining information on covariates was not described but given the description of other methods the use of questionnaires or interviews is likely. Braun et al. 2017 (HEROID: 6749104) also reports that clinical information was abstracted from medical records.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPEs were examined, including the urinary metabolites BDCIPP, DNBP, and DPHP. Models accounted for each OPE metabolite separately.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This cohort study analyzed associations between gestational exposure to BCEP with infant anthropometric measurements. The use of a cohort study design was appropriate for the research question given that gestational exposure was measured and participants were prospectively followed for pregnancy outcomes. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE) to examine whether OPE concentrations in different windows was related to pregnancy outcomes, as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth.
	Metric 13:	Statistical Power	Medium	The analysis sample included 340 mother-infant dyads, which is likely a large enough sample size to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and conceptually reproducible.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for participant characteristics as well as for all outcomes by participant characteristics. All models included adjustment for covariates, as well as stratified analyses for potential mediators, such as infant sex. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE), as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth. Cox proportional hazards regression analyses with gestational age as the underlying time scale were used to estimate hazard ratios (HRs) assessing the occurrence of preterm birth according to OPE metabolite concentrations.BCEP was analyzed using log10 transformed continuous variables or tertiles.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP is derived from parent chemical TCEP and is an appropriate biomarker of exposure.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure: 0.1 μg/L.
	Metric 19:	Biomarker Stability	Medium	Samples were stored at or below -20 °C until further analysis. They were shipped overnight on dry ice for quantification of OPEs. There is no data on stability, but no evidence of degradation and quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	Maternal urine samples were collected in polypropylene specimen cups. There is no information included about contamination.However, the CDC laboratory that conducted the analysis is certified by CLIA and quality control methods are detailed. There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	High	The laboratory methods described appeared to be appropriate. Samples were analyzed using reversed-phase high-performance liquid chromatography, and detection by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	BCEP concentrations were adjusted for specific gravity: specific gravity standardized concentrations were calculated to account for hydration status during pregnancy. However, study only provides results for matrix-adjusted concentrations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric	Rating	Comments
Additional Comments:	This cohort (340 mother-infant dyads) study in the Greater Cincinnati Area (Ohio, US) analyzed the relationship between gestational exposure to bis(2-chloroethyl) phosphate (BCEP) and gestational age at birth and newborn anthropometric measures.The study found an association between BCEP and several outcomes: increased BCEP concentrations in maternal urine at 16 weeks or 26 weeks of gestation were associated with longer gestation and reduced risk of preterm birth. BCEP was also negatively associated with weight and length z-scores at birth in females. However, the associations were not statistically significant after adjustment for multiple comparisons. Strengths of this analysis include exposure measurements at two time points during pregnancy, to examine windows of susceptibility (although both occurred in the second trimester), as well as extensive data on covariates. Another limitation is short-lived BCEP in urine samples., although this is a well-established biomarker.		

Overall Quality Determination **High**

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	This prospective birth cohort study used data from the Health Outcomes and Measures of the Environment (HOME) study, which was conducted in the Greater Cincinnati Metropolitan Area, Ohio. Pregnant women were recruited from March 2003-January 2006 Participants were recruited "using the medical scheduling systems of nine prenatal practices affiliated with three hospitals," with eligibility determined "using clinic records and phone interviews with women" (Braun et al., 2017; HEROID: 6749104). Inclusion criteria were: 1) age >18 years, 2) 13–19 weeks pregnancy, 3) residing in a home built in or before 1978, 5) fluent in English, 6) planning to live in the study area for the next year, 7) planning to continue prenatal care and deliver at the participating hospitals. Exclusion criteria were: living in a mobile or trailer home, were on medications for thyroid or seizure disorders, or diagnosed with bipolar disorder, schizophrenia, diabetes, or cancer requiring radiation or chemotherapy. The participation rate at each step was reported in Braun et al. 2017. From a sampling frame of 8878 women, 1263 were eligible and 340 women were included in the final analytic sample. The study compares those included in the original cohort (n=389) and those finally included in the present analysis (n=340) and reports that maternal characteristics were comparable between the two groups. There is no evidence to suggest substantial selection bias and key elements of the study design were reported.
	Metric 2:	Attrition	Medium	Details on attrition are available in Braun et al. 2017 (HEROID: 6749104). Of 468 women enrolled, 67 dropped out before delivery due to concerns about the time commitment, having family members unwilling to participate, or loss of contact. 389 women delivered singleton births, but only 340 were included in the final analysis for having no congenital malformations, having one spot urine sample quantified for OPE metabolites, and having neonatal anthropometry abstracted from medical records. While it is not clear why these measures were not available for all 389 women, exclusion from the analysis sample is appropriate and there is no evidence that any attrition in the study was related to exposure and outcome.
	Metric 3:	Comparison Group	High	The study reports on demographic characteristics and reports whether differences in characteristics were statistically significant by exposure and outcome variables. While significant differences were reported, these factors were considered as potential covariates in statistical analysis. There is no evidence to suggest that participants were selected differently by outcome or exposure group.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was assessed using an established biomarker of TCEP exposure (the urinary metabolite BCEP) in urine samples collected from mothers twice during pregnancy at an average of 16 $\pm$ 2 and 26 $\pm$ 3 weeks of gestation. BCEP was quantified using high-performance liquid chromatography with tandem mass spectrometry. LOD information is reported. Overall, exposure to TCEP was well-conducted and the use of two measurements increases the certainty in exposure concentrations across pregnancy.
	Metric 5:	Exposure Levels	Medium	The range of exposure is adequate, with no evidence of insufficient variability to estimate associations. BCEP was characterized into tertiles and also analyzed as continuous variable.
	Metric 6:	Temporality	Medium	The study includes measurements of urinary metabolites at two time points: 16 and 26 weeks, for pregnancy outcomes (gestational age and neonatal anthropometric measures). Temporality is therefore clearly established, with exposure preceding outcome. While the exposure is measured during pregnancy both time points captured are during the second trimester, meaning the first trimester, a relevant time period and critical window for fetal development was missed. Additionally, it is unclear whether the second trimester is the most etiologically relevant time window for developmental effects of BCEP.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Medical charts were used to obtain gestational age and infant anthropometric parameters at birth (birth weight (g), length (cm), and head circumference (cm). Standardized birth weight, length, and head circumference z-scores were calculated using values from the 2010 Olsen growth charts. Ponderal index (PI) was calculated as 100x (weight/length). One limitation is that gestational age was estimated by last menstrual period (LMP) for 330 participants, but by other measures (ultrasound (n = 7) or Ballard scores (n = 2) for very limited cases). While it is a widely used method to estimate gestational age, LMP has limitations as it relies on participant recall and also assumes regular 28-day menstrual cycles with conception occurring on the 14th day.Preterm birth was defined as birth prior to 37 completed weeks of gestation.
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed and reported as described in the study aims. Effect estimates are reported with confidence intervals. While the number of exposed/unexposed are not reported in each analysis, the number of observations used are reported consistently for each analysis.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were identified a priori through use of a directed acyclic graph (DAG). Considered covariates included household income, infant sex, marital status, maternal age, maternal blood lead, maternal serum cotinine, maternal education, maternal race, parity, pre-pregnancy BMI. Effect modification by sex was also explored. Key covariates were thus considered, and potential residual confounding is unlikely to have a significant impact on effect estimates.
	Metric 10:	Covariate Characterization	Medium	The exact methodology for obtaining information on covariates was not described but given the description of other methods the use of questionnaires or interviews is likely. Braun et al. 2017 (HEROID: 6749104) also reports that clinical information was abstracted from medical records.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPEs were examined, including the urinary metabolites BDCIPP, DNBP, and DPHP. Models accounted for each OPE metabolite separately.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This cohort study analyzed associations between gestational exposure to BCEP with infant anthropometric measurements. The use of a cohort study design was appropriate for the research question given that gestational exposure was measured and participants were prospectively followed for pregnancy outcomes. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE) to examine whether OPE concentrations in different windows was related to pregnancy outcomes, as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth.
	Metric 13:	Statistical Power	Medium	The analysis sample included 340 mother-infant dyads, which is likely a large enough sample size to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and conceptually reproducible.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for participant characteristics as well as for all outcomes by participant characteristics. All models included adjustment for covariates, as well as stratified analyses for potential mediators, such as infant sex. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE), as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth. Cox proportional hazards regression analyses with gestational age as the underlying time scale were used to estimate hazard ratios (HRs) assessing the occurrence of preterm birth according to OPE metabolite concentrations.BCEP was analyzed using log10 transformed continuous variables or tertiles.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP is derived from parent chemical TCEP and is an appropriate biomarker of exposure.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure: 0.1 μg/L.
	Metric 19:	Biomarker Stability	Medium	Samples were stored at or below -20 °C until further analysis. They were shipped overnight on dry ice for quantification of OPEs. There is no data on stability, but no evidence of degradation and quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	Maternal urine samples were collected in polypropylene specimen cups. There is no information included about contamination.However, the CDC laboratory that conducted the analysis is certified by CLIA and quality control methods are detailed. There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	High	The laboratory methods described appeared to be appropriate. Samples were analyzed using reversed-phase high-performance liquid chromatography, and detection by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	BCEP concentrations were adjusted for specific gravity: specific gravity standardized concentrations were calculated to account for hydration status during pregnancy. However, study only provides results for matrix-adjusted concentrations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric	Rating	Comments
Additional Comments:	This cohort (340 mother-infant dyads) study in the Greater Cincinnati Area (Ohio, US) analyzed the relationship between gestational exposure to bis(2-chloroethyl) phosphate (BCEP) and gestational age at birth and newborn anthropometric measures.The study found an association between BCEP and several outcomes: increased BCEP concentrations in maternal urine at 16 weeks or 26 weeks of gestation were associated with longer gestation and reduced risk of preterm birth. BCEP was also negatively associated with weight and length z-scores at birth in females. However, the associations were not statistically significant after adjustment for multiple comparisons. Strengths of this analysis include exposure measurements at two time points during pregnancy, to examine windows of susceptibility (although both occurred in the second trimester), as well as extensive data on covariates. Another limitation is short-lived BCEP in urine samples., although this is a well-established biomarker.		

Overall Quality Determination **High**

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	This prospective birth cohort study used data from the Health Outcomes and Measures of the Environment (HOME) study, which was conducted in the Greater Cincinnati Metropolitan Area, Ohio. Pregnant women were recruited from March 2003-January 2006 Participants were recruited "using the medical scheduling systems of nine prenatal practices affiliated with three hospitals," with eligibility determined "using clinic records and phone interviews with women" (Braun et al., 2017; HEROID: 6749104). Inclusion criteria were: 1) age >18 years, 2) 13–19 weeks pregnancy, 3) residing in a home built in or before 1978, 5) fluent in English, 6) planning to live in the study area for the next year, 7) planning to continue prenatal care and deliver at the participating hospitals. Exclusion criteria were: living in a mobile or trailer home, were on medications for thyroid or seizure disorders, or diagnosed with bipolar disorder, schizophrenia, diabetes, or cancer requiring radiation or chemotherapy. The participation rate at each step was reported in Braun et al. 2017. From a sampling frame of 8878 women, 1263 were eligible and 340 women were included in the final analytic sample. The study compares those included in the original cohort (n=389) and those finally included in the present analysis (n=340) and reports that maternal characteristics were comparable between the two groups. There is no evidence to suggest substantial selection bias and key elements of the study design were reported.
	Metric 2:	Attrition	Medium	Details on attrition are available in Braun et al. 2017 (HEROID: 6749104). Of 468 women enrolled, 67 dropped out before delivery due to concerns about the time commitment, having family members unwilling to participate, or loss of contact. 389 women delivered singleton births, but only 340 were included in the final analysis for having no congenital malformations, having one spot urine sample quantified for OPE metabolites, and having neonatal anthropometry abstracted from medical records. While it is not clear why these measures were not available for all 389 women, exclusion from the analysis sample is appropriate and there is no evidence that any attrition in the study was related to exposure and outcome.
	Metric 3:	Comparison Group	High	The study reports on demographic characteristics and reports whether differences in characteristics were statistically significant by exposure and outcome variables. While significant differences were reported, these factors were considered as potential covariates in statistical analysis. There is no evidence to suggest that participants were selected differently by outcome or exposure group.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was assessed using an established biomarker of TCEP exposure (the urinary metabolite BCEP) in urine samples collected from mothers twice during pregnancy at an average of 16 $\pm$ 2 and 26 $\pm$ 3 weeks of gestation. BCEP was quantified using high-performance liquid chromatography with tandem mass spectrometry. LOD information is reported. Overall, exposure to TCEP was well-conducted and the use of two measurements increases the certainty in exposure concentrations across pregnancy.
	Metric 5:	Exposure Levels	Medium	The range of exposure is adequate, with no evidence of insufficient variability to estimate associations. BCEP was characterized into tertiles and also analyzed as continuous variable.
	Metric 6:	Temporality	Medium	The study includes measurements of urinary metabolites at two time points: 16 and 26 weeks, for pregnancy outcomes (gestational age and neonatal anthropometric measures). Temporality is therefore clearly established, with exposure preceding outcome. While the exposure is measured during pregnancy both time points captured are during the second trimester, meaning the first trimester, a relevant time period and critical window for fetal development was missed. Additionally, it is unclear whether the second trimester is the most etiologically relevant time window for developmental effects of BCEP.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Medical charts were used to obtain gestational age and infant anthropometric parameters at birth (birth weight (g), length (cm), and head circumference (cm). Standardized birth weight, length, and head circumference z-scores were calculated using values from the 2010 Olsen growth charts. Ponderal index (PI) was calculated as 100x (weight/length). One limitation is that gestational age was estimated by last menstrual period (LMP) for 330 participants, but by other measures (ultrasound (n = 7) or Ballard scores (n = 2) for very limited cases). While it is a widely used method to estimate gestational age, LMP has limitations as it relies on participant recall and also assumes regular 28-day menstrual cycles with conception occurring on the 14th day.Preterm birth was defined as birth prior to 37 completed weeks of gestation.
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed and reported as described in the study aims. Effect estimates are reported with confidence intervals. While the number of exposed/unexposed are not reported in each analysis, the number of observations used are reported consistently for each analysis.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were identified a priori through use of a directed acyclic graph (DAG). Considered covariates included household income, infant sex, marital status, maternal age, maternal blood lead, maternal serum cotinine, maternal education, maternal race, parity, pre-pregnancy BMI. Effect modification by sex was also explored. Key covariates were thus considered, and potential residual confounding is unlikely to have a significant impact on effect estimates.
	Metric 10:	Covariate Characterization	Medium	The exact methodology for obtaining information on covariates was not described but given the description of other methods the use of questionnaires or interviews is likely. Braun et al. 2017 (HEROID: 6749104) also reports that clinical information was abstracted from medical records.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPEs were examined, including the urinary metabolites BDCIPP, DNBP, and DPHP. Models accounted for each OPE metabolite separately.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This cohort study analyzed associations between gestational exposure to BCEP with infant anthropometric measurements. The use of a cohort study design was appropriate for the research question given that gestational exposure was measured and participants were prospectively followed for pregnancy outcomes. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE) to examine whether OPE concentrations in different windows was related to pregnancy outcomes, as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth.
	Metric 13:	Statistical Power	Medium	The analysis sample included 340 mother-infant dyads, which is likely a large enough sample size to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and conceptually reproducible.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for participant characteristics as well as for all outcomes by participant characteristics. All models included adjustment for covariates, as well as stratified analyses for potential mediators, such as infant sex. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE), as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth. Cox proportional hazards regression analyses with gestational age as the underlying time scale were used to estimate hazard ratios (HRs) assessing the occurrence of preterm birth according to OPE metabolite concentrations.BCEP was analyzed using log10 transformed continuous variables or tertiles.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP is derived from parent chemical TCEP and is an appropriate biomarker of exposure.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure: 0.1 μg/L.
	Metric 19:	Biomarker Stability	Medium	Samples were stored at or below -20 °C until further analysis. They were shipped overnight on dry ice for quantification of OPEs. There is no data on stability, but no evidence of degradation and quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	Maternal urine samples were collected in polypropylene specimen cups. There is no information included about contamination.However, the CDC laboratory that conducted the analysis is certified by CLIA and quality control methods are detailed. There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	High	The laboratory methods described appeared to be appropriate. Samples were analyzed using reversed-phase high-performance liquid chromatography, and detection by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	BCEP concentrations were adjusted for specific gravity: specific gravity standardized concentrations were calculated to account for hydration status during pregnancy. However, study only provides results for matrix-adjusted concentrations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric	Rating	Comments
Additional Comments:	This cohort (340 mother-infant dyads) study in the Greater Cincinnati Area (Ohio, US) analyzed the relationship between gestational exposure to bis(2-chloroethyl) phosphate (BCEP) and gestational age at birth and newborn anthropometric measures.The study found an association between BCEP and several outcomes: increased BCEP concentrations in maternal urine at 16 weeks or 26 weeks of gestation were associated with longer gestation and reduced risk of preterm birth. BCEP was also negatively associated with weight and length z-scores at birth in females. However, the associations were not statistically significant after adjustment for multiple comparisons. Strengths of this analysis include exposure measurements at two time points during pregnancy, to examine windows of susceptibility (although both occurred in the second trimester), as well as extensive data on covariates. Another limitation is short-lived BCEP in urine samples., although this is a well-established biomarker.		

Overall Quality Determination **High**

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	This prospective birth cohort study used data from the Health Outcomes and Measures of the Environment (HOME) study, which was conducted in the Greater Cincinnati Metropolitan Area, Ohio. Pregnant women were recruited from March 2003-January 2006 Participants were recruited "using the medical scheduling systems of nine prenatal practices affiliated with three hospitals," with eligibility determined "using clinic records and phone interviews with women" (Braun et al., 2017; HEROID: 6749104). Inclusion criteria were: 1) age >18 years, 2) 13–19 weeks pregnancy, 3) residing in a home built in or before 1978, 5) fluent in English, 6) planning to live in the study area for the next year, 7) planning to continue prenatal care and deliver at the participating hospitals. Exclusion criteria were: living in a mobile or trailer home, were on medications for thyroid or seizure disorders, or diagnosed with bipolar disorder, schizophrenia, diabetes, or cancer requiring radiation or chemotherapy. The participation rate at each step was reported in Braun et al. 2017. From a sampling frame of 8878 women, 1263 were eligible and 340 women were included in the final analytic sample. The study compares those included in the original cohort (n=389) and those finally included in the present analysis (n=340) and reports that maternal characteristics were comparable between the two groups. There is no evidence to suggest substantial selection bias and key elements of the study design were reported.
	Metric 2:	Attrition	Medium	Details on attrition are available in Braun et al. 2017 (HEROID: 6749104). Of 468 women enrolled, 67 dropped out before delivery due to concerns about the time commitment, having family members unwilling to participate, or loss of contact. 389 women delivered singleton births, but only 340 were included in the final analysis for having no congenital malformations, having one spot urine sample quantified for OPE metabolites, and having neonatal anthropometry abstracted from medical records. While it is not clear why these measures were not available for all 389 women, exclusion from the analysis sample is appropriate and there is no evidence that any attrition in the study was related to exposure and outcome.
	Metric 3:	Comparison Group	High	The study reports on demographic characteristics and reports whether differences in characteristics were statistically significant by exposure and outcome variables. While significant differences were reported, these factors were considered as potential covariates in statistical analysis. There is no evidence to suggest that participants were selected differently by outcome or exposure group.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was assessed using an established biomarker of TCEP exposure (the urinary metabolite BCEP) in urine samples collected from mothers twice during pregnancy at an average of 16 $\pm$ 2 and 26 $\pm$ 3 weeks of gestation. BCEP was quantified using high-performance liquid chromatography with tandem mass spectrometry. LOD information is reported. Overall, exposure to TCEP was well-conducted and the use of two measurements increases the certainty in exposure concentrations across pregnancy.
	Metric 5:	Exposure Levels	Medium	The range of exposure is adequate, with no evidence of insufficient variability to estimate associations. BCEP was characterized into tertiles and also analyzed as continuous variable.
	Metric 6:	Temporality	Medium	The study includes measurements of urinary metabolites at two time points: 16 and 26 weeks, for pregnancy outcomes (gestational age and neonatal anthropometric measures). Temporality is therefore clearly established, with exposure preceding outcome. While the exposure is measured during pregnancy both time points captured are during the second trimester, meaning the first trimester, a relevant time period and critical window for fetal development was missed. Additionally, it is unclear whether the second trimester is the most etiologically relevant time window for developmental effects of BCEP.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Medical charts were used to obtain gestational age and infant anthropometric parameters at birth (birth weight (g), length (cm), and head circumference (cm). Standardized birth weight, length, and head circumference z-scores were calculated using values from the 2010 Olsen growth charts. Ponderal index (PI) was calculated as 100x (weight/length). One limitation is that gestational age was estimated by last menstrual period (LMP) for 330 participants, but by other measures (ultrasound (n = 7) or Ballard scores (n = 2) for very limited cases). While it is a widely used method to estimate gestational age, LMP has limitations as it relies on participant recall and also assumes regular 28-day menstrual cycles with conception occurring on the 14th day.Preterm birth was defined as birth prior to 37 completed weeks of gestation.
	Metric 8:	Reporting Bias	Medium	All measured outcomes were analyzed and reported as described in the study aims. Effect estimates are reported with confidence intervals. While the number of exposed/unexposed are not reported in each analysis, the number of observations used are reported consistently for each analysis.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were identified a priori through use of a directed acyclic graph (DAG). Considered covariates included household income, infant sex, marital status, maternal age, maternal blood lead, maternal serum cotinine, maternal education, maternal race, parity, pre-pregnancy BMI. Effect modification by sex was also explored. Key covariates were thus considered, and potential residual confounding is unlikely to have a significant impact on effect estimates.
	Metric 10:	Covariate Characterization	Medium	The exact methodology for obtaining information on covariates was not described but given the description of other methods the use of questionnaires or interviews is likely. Braun et al. 2017 (HEROID: 6749104) also reports that clinical information was abstracted from medical records.
	Metric 11:	Co-exposure Counfounding	Medium	Other OPEs were examined, including the urinary metabolites BDCIPP, DNBP, and DPHP. Models accounted for each OPE metabolite separately.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	This cohort study analyzed associations between gestational exposure to BCEP with infant anthropometric measurements. The use of a cohort study design was appropriate for the research question given that gestational exposure was measured and participants were prospectively followed for pregnancy outcomes. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE) to examine whether OPE concentrations in different windows was related to pregnancy outcomes, as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth.
	Metric 13:	Statistical Power	Medium	The analysis sample included 340 mother-infant dyads, which is likely a large enough sample size to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	Analysis methods were clearly described and conceptually reproducible.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Descriptive data were provided for participant characteristics as well as for all outcomes by participant characteristics. All models included adjustment for covariates, as well as stratified analyses for potential mediators, such as infant sex. Appropriate statistical analyses were used to account for exposure measurement twice during pregnancy, including multiple informant models (two exposure windows as informants) using non-standard generalized estimating equations (GEE), as well as using a modified Poisson regression with robust error variance combined with multiple informant models to examine period-specific risk ratios (RR) of individual OPE metabolites with preterm birth. Cox proportional hazards regression analyses with gestational age as the underlying time scale were used to estimate hazard ratios (HRs) assessing the occurrence of preterm birth according to OPE metabolite concentrations.BCEP was analyzed using log10 transformed continuous variables or tertiles.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP is derived from parent chemical TCEP and is an appropriate biomarker of exposure.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The method detection limit was provided for TCEP exposure: 0.1 μg/L.
	Metric 19:	Biomarker Stability	Medium	Samples were stored at or below -20 °C until further analysis. They were shipped overnight on dry ice for quantification of OPEs. There is no data on stability, but no evidence of degradation and quality control methods are detailed.
	Metric 20:	Sample Contamination	Medium	Maternal urine samples were collected in polypropylene specimen cups. There is no information included about contamination.However, the CDC laboratory that conducted the analysis is certified by CLIA and quality control methods are detailed. There was no evidence to suggest contamination problems.
	Metric 21:	Method Requirements	High	The laboratory methods described appeared to be appropriate. Samples were analyzed using reversed-phase high-performance liquid chromatography, and detection by isotope dilution-electrospray ionization tandem mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	BCEP concentrations were adjusted for specific gravity: specific gravity standardized concentrations were calculated to account for hydration status during pregnancy. However, study only provides results for matrix-adjusted concentrations.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Yang, W. L., Braun, J. M., Vuong, A. M., Percy, Z., Xu, Y. Y., Xie, C. C., Deka, R., Calafat, A. M., Ospina, M., Burris, H. H., Yolton, K., Cecil, K. M., Lanphear, B. P., Chen, A. M. (2022). Associations of gestational exposure to organophosphate esters with gestational age and neonatal anthropometric measures: The HOME study*. Environmental Pollution 316(Part 1):120516.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Gestational age at birth, newborn weight, length, head circumference, preterm birth, ponderal index
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: bis (2-chloroethyl) phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577121

Domain	Metric	Rating	Comments
Additional Comments:	This cohort (340 mother-infant dyads) study in the Greater Cincinnati Area (Ohio, US) analyzed the relationship between gestational exposure to bis(2-chloroethyl) phosphate (BCEP) and gestational age at birth and newborn anthropometric measures.The study found an association between BCEP and several outcomes: increased BCEP concentrations in maternal urine at 16 weeks or 26 weeks of gestation were associated with longer gestation and reduced risk of preterm birth. BCEP was also negatively associated with weight and length z-scores at birth in females. However, the associations were not statistically significant after adjustment for multiple comparisons. Strengths of this analysis include exposure measurements at two time points during pregnancy, to examine windows of susceptibility (although both occurred in the second trimester), as well as extensive data on covariates. Another limitation is short-lived BCEP in urine samples., although this is a well-established biomarker.		

Overall Quality Determination **High**

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants in this study were all recruited from cohorts that were enrolled in the "Environmental influences on Child Health Outcomes" (ECHO) program at the NIH, where existing cohorts enrolled their participants into the ECHO program set of cohorts. All cohorts included were birth cohorts with urine samples from pregnant mothers and measures of birth outcomes. Individual cohorts followed their own protocols prior to 2019, but in late 2019 ECHO instituted a common protocol for cohorts to follow. Data for this study was collected prior to initiation of the common protocol. In total 16 separate cohorts contributed data to this study, with data collected and submitted to ECHO prior to March 2022. Study details are available for each cohort, including the number of participants included, a description of the sample, eligibility criteria, and their strategy for obtaining birth outcome information. Of the ECHO cohort participants eligible based on the availability of maternal urine samples for organophosphate ester (OPE) quantification (n=12,873), n=7,048 (54.8%) participants had urinary OPE and dilution data. The number of urine samples included and excluded from each cohort was provided. Participants were excluded if they had no information on children (n=82), had multiple births (n=10), or had missing outcome data (n=309). One additional child born with a gestational age of >42 weeks was also excluded, as the formula used for calculating birth weight for gestational age z-scores did not accommodate gestations of that duration. In total, the final sample was n=6,646 mother-child dyads. The study compared the characteristics of participants included to those excluded; while no statistical analysis was done across these two groups, there do not appear to be significant differences in characteristics (e.g. socio-demographics, BMI, tobacco use) of included vs. excluded participants. Overall, key study information for the overall ECHO program and for individual cohorts was well-reported and indicated that significant selection bias was unlikely.
	Metric 2:	Attrition	Medium	The study indicates that 309 out of 7,048 mother-child dyads with urinary OPE measures had missing birth outcome data (gestational age, birth weight, biological sex at birth) and were thus excluded from the study. Another 82 pregnancies had no available child information, and 11 additional children were ineligible (multiple births, length of gestation). The study does not provide details as to why some dyads were missing outcome data, but there is no evidence that any cause would be related to their outcome status.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 3:	Comparison Group	Medium	Since this study was a pooled analyses of multiple birth cohorts, participants were inherently drawn from different background populations and locations across the United States. Key demographic and study design characteristics that may have differed across cohorts were considered as potential covariates (e.g., maternal race/ethnicity, education, age, marital status, along with timing, season, and year of sample collection). Clustering by cohort was addressed in the statistical model. The authors did not discuss adjusting for region or examining potential modification by cohort or region. Instead, to examine the potential influence of heterogeneity across cohorts, the authors included a sensitivity analysis using a "leave-one-cohort-out" approach to assess the influence of each cohort. While there was generally little or no impact of leaving out any one cohort, for the TCEP metabolite BCEP (alternative acronym BCETP) omitting one of the two largest cohorts (CANDLE, predominantly African American mothers from Memphis, TN) strengthened a positive association between low concentrations and birth weight z-scores, which became statistically significant. While there was a change in magnitude and statistical significance, there was no change in direction, and no evidence of bias due to cohort heterogeneity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	In all participants, TCEP was measured via the urinary metabolite BCEP. Urine samples were collected between 2007 and 2020. Measures used a single spot or first morning urine samples collected from each participant primarily during the second (44.1%) and third (55.5%) trimesters of pregnancy. BCEP is an appropriate biomarker of exposure to TCEP. BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control measures are well-detailed. The LOD was specified to be 0.02 ng/mL and the percentage of samples above the LOD was specified to be 69%. Machine readings were used for values below LOD. Urinary dilution was addressed using creatinine or specific gravity, depending on the cohort. Exposure measures at only a single time point may not reflect the full scope of TCEP exposure across pregnancy due to variability. In addition, other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372); it is uncertain whether or to what extent excluding this factor may introduce error. Though some error is likely, there was no evidence of important error or bias in exposure estimates.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	While the 25th percentile of exposure was below the limit of detection, the median value was 0.52 ng/mL and the 75th and 95th percentiles were 1.58 and 8.22 ng/mL, respectively, which should allow for sufficient contrast compared to participants with no detectable BCEP exposure. In statistical analyses, BCEP was categorized into three groups: undetected, and then low or high detectable exposure depending on whether measurements fell above or below the median of dilution-adjusted samples above the LOD.
	Metric 6:	Temporality	Medium	Temporality is established as exposure was measured during the second or third trimester of pregnancy for outcomes that were measured at birth. However, it is uncertain whether exposures during the second or third trimester reflect the etiologically relevant time window for birth outcome effects in relation to TCEP exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes collected prior to 2019 were ascertained using cohort-specific protocols; the majority of cohorts used medical record abstraction for outcomes, with a small proportion of parent reporting. Children born in or after 2019 had outcomes ascertained following the universal ECHO protocol of assigning outcomes based on medical record abstraction.All instances of birth weight were measured using medical records. Gestational age was largely abstracted from or estimated using data from medical records; methods used by hospitals to estimate gestational age were not always specified. Some cohorts specified using 1st or 2nd trimester ultrasound to calculate gestational age, using last menstrual period in cases of missing data. While there was some variation in the estimation of gestational age, there was no evidence that this variation might significantly impact effect estimates. Gestational age was further categorized as preterm (<37 weeks), early term (37-38 weeks), full term (39-40 weeks), and late/post-term (41-42 weeks). Sex-specific birth weight for gestational age z-scores were calculated based on a United states reference population (Aris et al. 2019, not in HERO; PMID: 31201230). Birth weight for gestational age was also categorized as small for gestational age (<10th percentile) and large for gestational age (>90th percentile), as well as term low birth weight (birth weight of <2,500 g at >=37 weeks).
	Metric 8:	Reporting Bias	Medium	All specified outcomes and results that were outlined in the methods were reported in the results. Effect estimates are presented with confidence intervals. Tabulations of outcomes by level of exposure were not provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Potential covariates were selected using a DAG to identify ad include confounders and precision variables, and to exclude intermediates. Covariates included were maternal race/ethnicity, maternal age at delivery, maternal education, maternal marital status, maternal pre-pregnancy BMI, maternal smoking during pregnancy, parity, infant sex, time of urine sample, season of urine sample, and year of urine sample. A potential limitation is that the authors did not mention evaluating the influence of gestational age adjustment on associations with size at birth. Though gestational age might be an intermediate or collider, there was no direct evidence of important error or bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were collected separately by each cohort within the ECHO program and then harmonized by the ECHO Data Analysis Center. While it is not specified how each of these covariates were measured, there is no evidence that insensitive instruments/methods/measures were used.
	Metric 11:	Co-exposure Counfounding	Medium	The study considered co-exposure to other organophosphate esters, including TPHP, TBUP, TIBP, TDCPP, TBOEP, TCPP, TMPP, TEHP, and TPRP. These co-exposures were measured via urinary metabolites at the same time and through the same process as BCEP. Spearman correlations between BECP and these biomarkers were <0.20. These co-exposures were considered in sensitivity analyses by simultaneously adjusting for these OPEs, and results were similar.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between prenatal concentrations of TCEP metabolites in urine and birth outcomes. The use of linear and logistic regression for continuous and categorical outcomes respectively is appropriate to address the research question. Regressions were performed using generalized estimating equations accounting for clustering at the cohort level, which is appropriate given the fact that cohorts were combined into one group in this study. Potential non-linear associations were examined by categorizing BCEP into three groups of similar size.
	Metric 13:	Statistical Power	Medium	The number of participants in the study (n=6,646 mother-child dyads) is likely sufficiently large to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Risk estimates were calculated using linear regression models for continuous outcomes and multinomial logistic regression models for categorical outcomes. BCEP concentrations were categorized, with values below LOD as the referent, and a lower and higher exposure group defined by categorizing remaining samples at their median. Inclusion of covariates was dependent on a DAG created to exclude variables that may be causal intermediates. Covariates with <20% of missing values were addressed via multiple imputation by chained equations. Sensitivity analyses are detailed, including sex stratification, a leave-one-cohort-out approach, and a co-pollutant model.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP are a valid and appropriate biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 0.02 ng/mL. BCEP was detected in 69% of samples. There was no evidence that the detection limit was inadequate.
	Metric 19:	Biomarker Stability	Medium	The study reports that urine samples were shipped on dry ice to the Human Health Exposure Analysis Resource (HHEAR) at the NYU Grossman School of Medicine. While no other specific storage information is provided, there is no evidence that inappropriate methods were used and the biomarker is not expected to have a high likelihood of stability.
	Metric 20:	Sample Contamination	Medium	No information is provided potential contamination.
	Metric 21:	Method Requirements	Medium	BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control (QC) measures are well-detailed. One QC measure included the use of blinded duplicate samples (n=127 samples for BCEP). For BCEP, 66% of duplicate pairs were both above LOD; the remaining pairs were not concordant with respect to detection. Details on the magnitude of disparities in estimated concentrations were not provided.
	Metric 22:	Matrix Adjustment	Medium	BCEP was adjusted for urinary dilution, measured as either specific gravity or creatinine depending on the source cohort, using published methods. Results were presented only for dilution-standardized exposure variables.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric	Rating	Comments
Additional Comments:	This prospective study used data from 16 cohorts participating in the ECHO (Environmental influences on Child Health Outcomes) program to estimate the association between maternal urinary organophosphate ester concentrations and two types of birth outcome: birth weight and duration of gestation. Spot urine concentrations of BCEP from the second or third trimester were used to estimate exposure. There were no major concerns for bias. There was some degree of uncertainty due to the use of only a single timepoint measurement of exposure and variability in ways in which gestational age was assessed. Birth weight was analyzed using birth weight for gestational age z-scores continuously or classified as small or large size for gestational age, but not in analyses that excluded gestational age adjustment. Among girls but not boys, compared to concentrations below detection, a high level of detectable BCEP was associated with significantly shorter mean gestational age and increased odds of preterm birth. In the population overall, a low concentration of detectable BCEP was associated with significantly reduced odds of late/post term birth; the association with high concentration was null. In addition, overall, a high concentration of detectable BCEP was associated with significantly lower odds of small for gestational age (SGA) birth. Both low and high detectable BCEP were associated with reduced odds of SGA among boys. Among girls, low-level detectable BCEP was associated with increased odds of large for gestational age (LGA) birth. Findings from this study suggest that effects of BCEP may be sex-specific and that BCEP may reduce the length of gestation particularly in girls but may not reduce birth weight relative to the duration of gestation.		

Overall Quality Determination **Medium**

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants in this study were all recruited from cohorts that were enrolled in the "Environmental influences on Child Health Outcomes" (ECHO) program at the NIH, where existing cohorts enrolled their participants into the ECHO program set of cohorts. All cohorts included were birth cohorts with urine samples from pregnant mothers and measures of birth outcomes. Individual cohorts followed their own protocols prior to 2019, but in late 2019 ECHO instituted a common protocol for cohorts to follow. Data for this study was collected prior to initiation of the common protocol. In total 16 separate cohorts contributed data to this study, with data collected and submitted to ECHO prior to March 2022. Study details are available for each cohort, including the number of participants included, a description of the sample, eligibility criteria, and their strategy for obtaining birth outcome information. Of the ECHO cohort participants eligible based on the availability of maternal urine samples for organophosphate ester (OPE) quantification (n=12,873), n=7,048 (54.8%) participants had urinary OPE and dilution data. The number of urine samples included and excluded from each cohort was provided. Participants were excluded if they had no information on children (n=82), had multiple births (n=10), or had missing outcome data (n=309). One additional child born with a gestational age of >42 weeks was also excluded, as the formula used for calculating birth weight for gestational age z-scores did not accommodate gestations of that duration. In total, the final sample was n=6,646 mother-child dyads. The study compared the characteristics of participants included to those excluded; while no statistical analysis was done across these two groups, there do not appear to be significant differences in characteristics (e.g. socio-demographics, BMI, tobacco use) of included vs. excluded participants. Overall, key study information for the overall ECHO program and for individual cohorts was well-reported and indicated that significant selection bias was unlikely.
	Metric 2:	Attrition	Medium	The study indicates that 309 out of 7,048 mother-child dyads with urinary OPE measures had missing birth outcome data (gestational age, birth weight, biological sex at birth) and were thus excluded from the study. Another 82 pregnancies had no available child information, and 11 additional children were ineligible (multiple births, length of gestation). The study does not provide details as to why some dyads were missing outcome data, but there is no evidence that any cause would be related to their outcome status.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 3:	Comparison Group	Medium	Since this study was a pooled analyses of multiple birth cohorts, participants were inherently drawn from different background populations and locations across the United States. Key demographic and study design characteristics that may have differed across cohorts were considered as potential covariates (e.g., maternal race/ethnicity, education, age, marital status, along with timing, season, and year of sample collection). Clustering by cohort was addressed in the statistical model. The authors did not discuss adjusting for region or examining potential modification by cohort or region. Instead, to examine the potential influence of heterogeneity across cohorts, the authors included a sensitivity analysis using a "leave-one-cohort-out" approach to assess the influence of each cohort. While there was generally little or no impact of leaving out any one cohort, for the TCEP metabolite BCEP (alternative acronym BCETP) omitting one of the two largest cohorts (CANDLE, predominantly African American mothers from Memphis, TN) strengthened a positive association between low concentrations and birth weight z-scores, which became statistically significant. While there was a change in magnitude and statistical significance, there was no change in direction, and no evidence of bias due to cohort heterogeneity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	In all participants, TCEP was measured via the urinary metabolite BCEP. Urine samples were collected between 2007 and 2020. Measures used a single spot or first morning urine samples collected from each participant primarily during the second (44.1%) and third (55.5%) trimesters of pregnancy. BCEP is an appropriate biomarker of exposure to TCEP. BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control measures are well-detailed. The LOD was specified to be 0.02 ng/mL and the percentage of samples above the LOD was specified to be 69%. Machine readings were used for values below LOD. Urinary dilution was addressed using creatinine or specific gravity, depending on the cohort. Exposure measures at only a single time point may not reflect the full scope of TCEP exposure across pregnancy due to variability. In addition, other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372); it is uncertain whether or to what extent excluding this factor may introduce error. Though some error is likely, there was no evidence of important error or bias in exposure estimates.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	While the 25th percentile of exposure was below the limit of detection, the median value was 0.52 ng/mL and the 75th and 95th percentiles were 1.58 and 8.22 ng/mL, respectively, which should allow for sufficient contrast compared to participants with no detectable BCEP exposure. In statistical analyses, BCEP was categorized into three groups: undetected, and then low or high detectable exposure depending on whether measurements fell above or below the median of dilution-adjusted samples above the LOD.
	Metric 6:	Temporality	Medium	Temporality is established as exposure was measured during the second or third trimester of pregnancy for outcomes that were measured at birth. However, it is uncertain whether exposures during the second or third trimester reflect the etiologically relevant time window for birth outcome effects in relation to TCEP exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes collected prior to 2019 were ascertained using cohort-specific protocols; the majority of cohorts used medical record abstraction for outcomes, with a small proportion of parent reporting. Children born in or after 2019 had outcomes ascertained following the universal ECHO protocol of assigning outcomes based on medical record abstraction.All instances of birth weight were measured using medical records. Gestational age was largely abstracted from or estimated using data from medical records; methods used by hospitals to estimate gestational age were not always specified. Some cohorts specified using 1st or 2nd trimester ultrasound to calculate gestational age, using last menstrual period in cases of missing data. While there was some variation in the estimation of gestational age, there was no evidence that this variation might significantly impact effect estimates. Gestational age was further categorized as preterm (<37 weeks), early term (37-38 weeks), full term (39-40 weeks), and late/post-term (41-42 weeks). Sex-specific birth weight for gestational age z-scores were calculated based on a United states reference population (Aris et al. 2019, not in HERO; PMID: 31201230). Birth weight for gestational age was also categorized as small for gestational age (<10th percentile) and large for gestational age (>90th percentile), as well as term low birth weight (birth weight of <2,500 g at >=37 weeks).
	Metric 8:	Reporting Bias	Medium	All specified outcomes and results that were outlined in the methods were reported in the results. Effect estimates are presented with confidence intervals. Tabulations of outcomes by level of exposure were not provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Potential covariates were selected using a DAG to identify ad include confounders and precision variables, and to exclude intermediates. Covariates included were maternal race/ethnicity, maternal age at delivery, maternal education, maternal marital status, maternal pre-pregnancy BMI, maternal smoking during pregnancy, parity, infant sex, time of urine sample, season of urine sample, and year of urine sample. A potential limitation is that the authors did not mention evaluating the influence of gestational age adjustment on associations with size at birth. Though gestational age might be an intermediate or collider, there was no direct evidence of important error or bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were collected separately by each cohort within the ECHO program and then harmonized by the ECHO Data Analysis Center. While it is not specified how each of these covariates were measured, there is no evidence that insensitive instruments/methods/measures were used.
	Metric 11:	Co-exposure Counfounding	Medium	The study considered co-exposure to other organophosphate esters, including TPHP, TBUP, TIBP, TDCPP, TBOEP, TCPP, TMPP, TEHP, and TPRP. These co-exposures were measured via urinary metabolites at the same time and through the same process as BCEP. Spearman correlations between BECP and these biomarkers were <0.20. These co-exposures were considered in sensitivity analyses by simultaneously adjusting for these OPEs, and results were similar.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between prenatal concentrations of TCEP metabolites in urine and birth outcomes. The use of linear and logistic regression for continuous and categorical outcomes respectively is appropriate to address the research question. Regressions were performed using generalized estimating equations accounting for clustering at the cohort level, which is appropriate given the fact that cohorts were combined into one group in this study. Potential non-linear associations were examined by categorizing BCEP into three groups of similar size.
	Metric 13:	Statistical Power	Medium	The number of participants in the study (n=6,646 mother-child dyads) is likely sufficiently large to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Risk estimates were calculated using linear regression models for continuous outcomes and multinomial logistic regression models for categorical outcomes. BCEP concentrations were categorized, with values below LOD as the referent, and a lower and higher exposure group defined by categorizing remaining samples at their median. Inclusion of covariates was dependent on a DAG created to exclude variables that may be causal intermediates. Covariates with <20% of missing values were addressed via multiple imputation by chained equations. Sensitivity analyses are detailed, including sex stratification, a leave-one-cohort-out approach, and a co-pollutant model.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP are a valid and appropriate biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 0.02 ng/mL. BCEP was detected in 69% of samples. There was no evidence that the detection limit was inadequate.
	Metric 19:	Biomarker Stability	Medium	The study reports that urine samples were shipped on dry ice to the Human Health Exposure Analysis Resource (HHEAR) at the NYU Grossman School of Medicine. While no other specific storage information is provided, there is no evidence that inappropriate methods were used and the biomarker is not expected to have a high likelihood of stability.
	Metric 20:	Sample Contamination	Medium	No information is provided potential contamination.
	Metric 21:	Method Requirements	Medium	BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control (QC) measures are well-detailed. One QC measure included the use of blinded duplicate samples (n=127 samples for BCEP). For BCEP, 66% of duplicate pairs were both above LOD; the remaining pairs were not concordant with respect to detection. Details on the magnitude of disparities in estimated concentrations were not provided.
	Metric 22:	Matrix Adjustment	Medium	BCEP was adjusted for urinary dilution, measured as either specific gravity or creatinine depending on the source cohort, using published methods. Results were presented only for dilution-standardized exposure variables.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric	Rating	Comments
Additional Comments:	This prospective study used data from 16 cohorts participating in the ECHO (Environmental influences on Child Health Outcomes) program to estimate the association between maternal urinary organophosphate ester concentrations and two types of birth outcome: birth weight and duration of gestation. Spot urine concentrations of BCEP from the second or third trimester were used to estimate exposure. There were no major concerns for bias. There was some degree of uncertainty due to the use of only a single timepoint measurement of exposure and variability in ways in which gestational age was assessed. Birth weight was analyzed using birth weight for gestational age z-scores continuously or classified as small or large size for gestational age, but not in analyses that excluded gestational age adjustment. Among girls but not boys, compared to concentrations below detection, a high level of detectable BCEP was associated with significantly shorter mean gestational age and increased odds of preterm birth. In the population overall, a low concentration of detectable BCEP was associated with significantly reduced odds of late/post term birth; the association with high concentration was null. In addition, overall, a high concentration of detectable BCEP was associated with significantly lower odds of small for gestational age (SGA) birth. Both low and high detectable BCEP were associated with reduced odds of SGA among boys. Among girls, low-level detectable BCEP was associated with increased odds of large for gestational age (LGA) birth. Findings from this study suggest that effects of BCEP may be sex-specific and that BCEP may reduce the length of gestation particularly in girls but may not reduce birth weight relative to the duration of gestation.		

Overall Quality Determination **Medium**

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants in this study were all recruited from cohorts that were enrolled in the "Environmental influences on Child Health Outcomes" (ECHO) program at the NIH, where existing cohorts enrolled their participants into the ECHO program set of cohorts. All cohorts included were birth cohorts with urine samples from pregnant mothers and measures of birth outcomes. Individual cohorts followed their own protocols prior to 2019, but in late 2019 ECHO instituted a common protocol for cohorts to follow. Data for this study was collected prior to initiation of the common protocol. In total 16 separate cohorts contributed data to this study, with data collected and submitted to ECHO prior to March 2022. Study details are available for each cohort, including the number of participants included, a description of the sample, eligibility criteria, and their strategy for obtaining birth outcome information. Of the ECHO cohort participants eligible based on the availability of maternal urine samples for organophosphate ester (OPE) quantification (n=12,873), n=7,048 (54.8%) participants had urinary OPE and dilution data. The number of urine samples included and excluded from each cohort was provided. Participants were excluded if they had no information on children (n=82), had multiple births (n=10), or had missing outcome data (n=309). One additional child born with a gestational age of >42 weeks was also excluded, as the formula used for calculating birth weight for gestational age z-scores did not accommodate gestations of that duration. In total, the final sample was n=6,646 mother-child dyads. The study compared the characteristics of participants included to those excluded; while no statistical analysis was done across these two groups, there do not appear to be significant differences in characteristics (e.g. socio-demographics, BMI, tobacco use) of included vs. excluded participants. Overall, key study information for the overall ECHO program and for individual cohorts was well-reported and indicated that significant selection bias was unlikely.
	Metric 2:	Attrition	Medium	The study indicates that 309 out of 7,048 mother-child dyads with urinary OPE measures had missing birth outcome data (gestational age, birth weight, biological sex at birth) and were thus excluded from the study. Another 82 pregnancies had no available child information, and 11 additional children were ineligible (multiple births, length of gestation). The study does not provide details as to why some dyads were missing outcome data, but there is no evidence that any cause would be related to their outcome status.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 3:	Comparison Group	Medium	Since this study was a pooled analyses of multiple birth cohorts, participants were inherently drawn from different background populations and locations across the United States. Key demographic and study design characteristics that may have differed across cohorts were considered as potential covariates (e.g., maternal race/ethnicity, education, age, marital status, along with timing, season, and year of sample collection). Clustering by cohort was addressed in the statistical model. The authors did not discuss adjusting for region or examining potential modification by cohort or region. Instead, to examine the potential influence of heterogeneity across cohorts, the authors included a sensitivity analysis using a "leave-one-cohort-out" approach to assess the influence of each cohort. While there was generally little or no impact of leaving out any one cohort, for the TCEP metabolite BCEP (alternative acronym BCETP) omitting one of the two largest cohorts (CANDLE, predominantly African American mothers from Memphis, TN) strengthened a positive association between low concentrations and birth weight z-scores, which became statistically significant. While there was a change in magnitude and statistical significance, there was no change in direction, and no evidence of bias due to cohort heterogeneity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	In all participants, TCEP was measured via the urinary metabolite BCEP. Urine samples were collected between 2007 and 2020. Measures used a single spot or first morning urine samples collected from each participant primarily during the second (44.1%) and third (55.5%) trimesters of pregnancy. BCEP is an appropriate biomarker of exposure to TCEP. BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control measures are well-detailed. The LOD was specified to be 0.02 ng/mL and the percentage of samples above the LOD was specified to be 69%. Machine readings were used for values below LOD. Urinary dilution was addressed using creatinine or specific gravity, depending on the cohort. Exposure measures at only a single time point may not reflect the full scope of TCEP exposure across pregnancy due to variability. In addition, other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372); it is uncertain whether or to what extent excluding this factor may introduce error. Though some error is likely, there was no evidence of important error or bias in exposure estimates.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	While the 25th percentile of exposure was below the limit of detection, the median value was 0.52 ng/mL and the 75th and 95th percentiles were 1.58 and 8.22 ng/mL, respectively, which should allow for sufficient contrast compared to participants with no detectable BCEP exposure. In statistical analyses, BCEP was categorized into three groups: undetected, and then low or high detectable exposure depending on whether measurements fell above or below the median of dilution-adjusted samples above the LOD.
	Metric 6:	Temporality	Medium	Temporality is established as exposure was measured during the second or third trimester of pregnancy for outcomes that were measured at birth. However, it is uncertain whether exposures during the second or third trimester reflect the etiologically relevant time window for birth outcome effects in relation to TCEP exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes collected prior to 2019 were ascertained using cohort-specific protocols; the majority of cohorts used medical record abstraction for outcomes, with a small proportion of parent reporting. Children born in or after 2019 had outcomes ascertained following the universal ECHO protocol of assigning outcomes based on medical record abstraction.All instances of birth weight were measured using medical records. Gestational age was largely abstracted from or estimated using data from medical records; methods used by hospitals to estimate gestational age were not always specified. Some cohorts specified using 1st or 2nd trimester ultrasound to calculate gestational age, using last menstrual period in cases of missing data. While there was some variation in the estimation of gestational age, there was no evidence that this variation might significantly impact effect estimates. Gestational age was further categorized as preterm (<37 weeks), early term (37-38 weeks), full term (39-40 weeks), and late/post-term (41-42 weeks). Sex-specific birth weight for gestational age z-scores were calculated based on a United states reference population (Aris et al. 2019, not in HERO; PMID: 31201230). Birth weight for gestational age was also categorized as small for gestational age (<10th percentile) and large for gestational age (>90th percentile), as well as term low birth weight (birth weight of <2,500 g at >=37 weeks).
	Metric 8:	Reporting Bias	Medium	All specified outcomes and results that were outlined in the methods were reported in the results. Effect estimates are presented with confidence intervals. Tabulations of outcomes by level of exposure were not provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Potential covariates were selected using a DAG to identify ad include confounders and precision variables, and to exclude intermediates. Covariates included were maternal race/ethnicity, maternal age at delivery, maternal education, maternal marital status, maternal pre-pregnancy BMI, maternal smoking during pregnancy, parity, infant sex, time of urine sample, season of urine sample, and year of urine sample. A potential limitation is that the authors did not mention evaluating the influence of gestational age adjustment on associations with size at birth. Though gestational age might be an intermediate or collider, there was no direct evidence of important error or bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were collected separately by each cohort within the ECHO program and then harmonized by the ECHO Data Analysis Center. While it is not specified how each of these covariates were measured, there is no evidence that insensitive instruments/methods/measures were used.
	Metric 11:	Co-exposure Counfounding	Medium	The study considered co-exposure to other organophosphate esters, including TPHP, TBUP, TIBP, TDCPP, TBOEP, TCPP, TMPP, TEHP, and TPRP. These co-exposures were measured via urinary metabolites at the same time and through the same process as BCEP. Spearman correlations between BECP and these biomarkers were <0.20. These co-exposures were considered in sensitivity analyses by simultaneously adjusting for these OPEs, and results were similar.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between prenatal concentrations of TCEP metabolites in urine and birth outcomes. The use of linear and logistic regression for continuous and categorical outcomes respectively is appropriate to address the research question. Regressions were performed using generalized estimating equations accounting for clustering at the cohort level, which is appropriate given the fact that cohorts were combined into one group in this study. Potential non-linear associations were examined by categorizing BCEP into three groups of similar size.
	Metric 13:	Statistical Power	Medium	The number of participants in the study (n=6,646 mother-child dyads) is likely sufficiently large to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Risk estimates were calculated using linear regression models for continuous outcomes and multinomial logistic regression models for categorical outcomes. BCEP concentrations were categorized, with values below LOD as the referent, and a lower and higher exposure group defined by categorizing remaining samples at their median. Inclusion of covariates was dependent on a DAG created to exclude variables that may be causal intermediates. Covariates with <20% of missing values were addressed via multiple imputation by chained equations. Sensitivity analyses are detailed, including sex stratification, a leave-one-cohort-out approach, and a co-pollutant model.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP are a valid and appropriate biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 0.02 ng/mL. BCEP was detected in 69% of samples. There was no evidence that the detection limit was inadequate.
	Metric 19:	Biomarker Stability	Medium	The study reports that urine samples were shipped on dry ice to the Human Health Exposure Analysis Resource (HHEAR) at the NYU Grossman School of Medicine. While no other specific storage information is provided, there is no evidence that inappropriate methods were used and the biomarker is not expected to have a high likelihood of stability.
	Metric 20:	Sample Contamination	Medium	No information is provided potential contamination.
	Metric 21:	Method Requirements	Medium	BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control (QC) measures are well-detailed. One QC measure included the use of blinded duplicate samples (n=127 samples for BCEP). For BCEP, 66% of duplicate pairs were both above LOD; the remaining pairs were not concordant with respect to detection. Details on the magnitude of disparities in estimated concentrations were not provided.
	Metric 22:	Matrix Adjustment	Medium	BCEP was adjusted for urinary dilution, measured as either specific gravity or creatinine depending on the source cohort, using published methods. Results were presented only for dilution-standardized exposure variables.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric	Rating	Comments
Additional Comments:	This prospective study used data from 16 cohorts participating in the ECHO (Environmental influences on Child Health Outcomes) program to estimate the association between maternal urinary organophosphate ester concentrations and two types of birth outcome: birth weight and duration of gestation. Spot urine concentrations of BCEP from the second or third trimester were used to estimate exposure. There were no major concerns for bias. There was some degree of uncertainty due to the use of only a single timepoint measurement of exposure and variability in ways in which gestational age was assessed. Birth weight was analyzed using birth weight for gestational age z-scores continuously or classified as small or large size for gestational age, but not in analyses that excluded gestational age adjustment. Among girls but not boys, compared to concentrations below detection, a high level of detectable BCEP was associated with significantly shorter mean gestational age and increased odds of preterm birth. In the population overall, a low concentration of detectable BCEP was associated with significantly reduced odds of late/post term birth; the association with high concentration was null. In addition, overall, a high concentration of detectable BCEP was associated with significantly lower odds of small for gestational age (SGA) birth. Both low and high detectable BCEP were associated with reduced odds of SGA among boys. Among girls, low-level detectable BCEP was associated with increased odds of large for gestational age (LGA) birth. Findings from this study suggest that effects of BCEP may be sex-specific and that BCEP may reduce the length of gestation particularly in girls but may not reduce birth weight relative to the duration of gestation.		

Overall Quality Determination **Medium**

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants in this study were all recruited from cohorts that were enrolled in the "Environmental influences on Child Health Outcomes" (ECHO) program at the NIH, where existing cohorts enrolled their participants into the ECHO program set of cohorts. All cohorts included were birth cohorts with urine samples from pregnant mothers and measures of birth outcomes. Individual cohorts followed their own protocols prior to 2019, but in late 2019 ECHO instituted a common protocol for cohorts to follow. Data for this study was collected prior to initiation of the common protocol. In total 16 separate cohorts contributed data to this study, with data collected and submitted to ECHO prior to March 2022. Study details are available for each cohort, including the number of participants included, a description of the sample, eligibility criteria, and their strategy for obtaining birth outcome information. Of the ECHO cohort participants eligible based on the availability of maternal urine samples for organophosphate ester (OPE) quantification (n=12,873), n=7,048 (54.8%) participants had urinary OPE and dilution data. The number of urine samples included and excluded from each cohort was provided. Participants were excluded if they had no information on children (n=82), had multiple births (n=10), or had missing outcome data (n=309). One additional child born with a gestational age of >42 weeks was also excluded, as the formula used for calculating birth weight for gestational age z-scores did not accommodate gestations of that duration. In total, the final sample was n=6,646 mother-child dyads. The study compared the characteristics of participants included to those excluded; while no statistical analysis was done across these two groups, there do not appear to be significant differences in characteristics (e.g. socio-demographics, BMI, tobacco use) of included vs. excluded participants. Overall, key study information for the overall ECHO program and for individual cohorts was well-reported and indicated that significant selection bias was unlikely.
	Metric 2:	Attrition	Medium	The study indicates that 309 out of 7,048 mother-child dyads with urinary OPE measures had missing birth outcome data (gestational age, birth weight, biological sex at birth) and were thus excluded from the study. Another 82 pregnancies had no available child information, and 11 additional children were ineligible (multiple births, length of gestation). The study does not provide details as to why some dyads were missing outcome data, but there is no evidence that any cause would be related to their outcome status.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 3:	Comparison Group	Medium	Since this study was a pooled analyses of multiple birth cohorts, participants were inherently drawn from different background populations and locations across the United States. Key demographic and study design characteristics that may have differed across cohorts were considered as potential covariates (e.g., maternal race/ethnicity, education, age, marital status, along with timing, season, and year of sample collection). Clustering by cohort was addressed in the statistical model. The authors did not discuss adjusting for region or examining potential modification by cohort or region. Instead, to examine the potential influence of heterogeneity across cohorts, the authors included a sensitivity analysis using a "leave-one-cohort-out" approach to assess the influence of each cohort. While there was generally little or no impact of leaving out any one cohort, for the TCEP metabolite BCEP (alternative acronym BCETP) omitting one of the two largest cohorts (CANDLE, predominantly African American mothers from Memphis, TN) strengthened a positive association between low concentrations and birth weight z-scores, which became statistically significant. While there was a change in magnitude and statistical significance, there was no change in direction, and no evidence of bias due to cohort heterogeneity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	In all participants, TCEP was measured via the urinary metabolite BCEP. Urine samples were collected between 2007 and 2020. Measures used a single spot or first morning urine samples collected from each participant primarily during the second (44.1%) and third (55.5%) trimesters of pregnancy. BCEP is an appropriate biomarker of exposure to TCEP. BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control measures are well-detailed. The LOD was specified to be 0.02 ng/mL and the percentage of samples above the LOD was specified to be 69%. Machine readings were used for values below LOD. Urinary dilution was addressed using creatinine or specific gravity, depending on the cohort. Exposure measures at only a single time point may not reflect the full scope of TCEP exposure across pregnancy due to variability. In addition, other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372); it is uncertain whether or to what extent excluding this factor may introduce error. Though some error is likely, there was no evidence of important error or bias in exposure estimates.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	While the 25th percentile of exposure was below the limit of detection, the median value was 0.52 ng/mL and the 75th and 95th percentiles were 1.58 and 8.22 ng/mL, respectively, which should allow for sufficient contrast compared to participants with no detectable BCEP exposure. In statistical analyses, BCEP was categorized into three groups: undetected, and then low or high detectable exposure depending on whether measurements fell above or below the median of dilution-adjusted samples above the LOD.
	Metric 6:	Temporality	Medium	Temporality is established as exposure was measured during the second or third trimester of pregnancy for outcomes that were measured at birth. However, it is uncertain whether exposures during the second or third trimester reflect the etiologically relevant time window for birth outcome effects in relation to TCEP exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes collected prior to 2019 were ascertained using cohort-specific protocols; the majority of cohorts used medical record abstraction for outcomes, with a small proportion of parent reporting. Children born in or after 2019 had outcomes ascertained following the universal ECHO protocol of assigning outcomes based on medical record abstraction.All instances of birth weight were measured using medical records. Gestational age was largely abstracted from or estimated using data from medical records; methods used by hospitals to estimate gestational age were not always specified. Some cohorts specified using 1st or 2nd trimester ultrasound to calculate gestational age, using last menstrual period in cases of missing data. While there was some variation in the estimation of gestational age, there was no evidence that this variation might significantly impact effect estimates. Gestational age was further categorized as preterm (<37 weeks), early term (37-38 weeks), full term (39-40 weeks), and late/post-term (41-42 weeks). Sex-specific birth weight for gestational age z-scores were calculated based on a United states reference population (Aris et al. 2019, not in HERO; PMID: 31201230). Birth weight for gestational age was also categorized as small for gestational age (<10th percentile) and large for gestational age (>90th percentile), as well as term low birth weight (birth weight of <2,500 g at >=37 weeks).
	Metric 8:	Reporting Bias	Medium	All specified outcomes and results that were outlined in the methods were reported in the results. Effect estimates are presented with confidence intervals. Tabulations of outcomes by level of exposure were not provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Potential covariates were selected using a DAG to identify ad include confounders and precision variables, and to exclude intermediates. Covariates included were maternal race/ethnicity, maternal age at delivery, maternal education, maternal marital status, maternal pre-pregnancy BMI, maternal smoking during pregnancy, parity, infant sex, time of urine sample, season of urine sample, and year of urine sample. A potential limitation is that the authors did not mention evaluating the influence of gestational age adjustment on associations with size at birth. Though gestational age might be an intermediate or collider, there was no direct evidence of important error or bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were collected separately by each cohort within the ECHO program and then harmonized by the ECHO Data Analysis Center. While it is not specified how each of these covariates were measured, there is no evidence that insensitive instruments/methods/measures were used.
	Metric 11:	Co-exposure Counfounding	Medium	The study considered co-exposure to other organophosphate esters, including TPHP, TBUP, TIBP, TDCPP, TBOEP, TCPP, TMPP, TEHP, and TPRP. These co-exposures were measured via urinary metabolites at the same time and through the same process as BCEP. Spearman correlations between BECP and these biomarkers were <0.20. These co-exposures were considered in sensitivity analyses by simultaneously adjusting for these OPEs, and results were similar.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between prenatal concentrations of TCEP metabolites in urine and birth outcomes. The use of linear and logistic regression for continuous and categorical outcomes respectively is appropriate to address the research question. Regressions were performed using generalized estimating equations accounting for clustering at the cohort level, which is appropriate given the fact that cohorts were combined into one group in this study. Potential non-linear associations were examined by categorizing BCEP into three groups of similar size.
	Metric 13:	Statistical Power	Medium	The number of participants in the study (n=6,646 mother-child dyads) is likely sufficiently large to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Risk estimates were calculated using linear regression models for continuous outcomes and multinomial logistic regression models for categorical outcomes. BCEP concentrations were categorized, with values below LOD as the referent, and a lower and higher exposure group defined by categorizing remaining samples at their median. Inclusion of covariates was dependent on a DAG created to exclude variables that may be causal intermediates. Covariates with <20% of missing values were addressed via multiple imputation by chained equations. Sensitivity analyses are detailed, including sex stratification, a leave-one-cohort-out approach, and a co-pollutant model.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP are a valid and appropriate biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 0.02 ng/mL. BCEP was detected in 69% of samples. There was no evidence that the detection limit was inadequate.
	Metric 19:	Biomarker Stability	Medium	The study reports that urine samples were shipped on dry ice to the Human Health Exposure Analysis Resource (HHEAR) at the NYU Grossman School of Medicine. While no other specific storage information is provided, there is no evidence that inappropriate methods were used and the biomarker is not expected to have a high likelihood of stability.
	Metric 20:	Sample Contamination	Medium	No information is provided potential contamination.
	Metric 21:	Method Requirements	Medium	BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control (QC) measures are well-detailed. One QC measure included the use of blinded duplicate samples (n=127 samples for BCEP). For BCEP, 66% of duplicate pairs were both above LOD; the remaining pairs were not concordant with respect to detection. Details on the magnitude of disparities in estimated concentrations were not provided.
	Metric 22:	Matrix Adjustment	Medium	BCEP was adjusted for urinary dilution, measured as either specific gravity or creatinine depending on the source cohort, using published methods. Results were presented only for dilution-standardized exposure variables.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric	Rating	Comments
Additional Comments:	This prospective study used data from 16 cohorts participating in the ECHO (Environmental influences on Child Health Outcomes) program to estimate the association between maternal urinary organophosphate ester concentrations and two types of birth outcome: birth weight and duration of gestation. Spot urine concentrations of BCEP from the second or third trimester were used to estimate exposure. There were no major concerns for bias. There was some degree of uncertainty due to the use of only a single timepoint measurement of exposure and variability in ways in which gestational age was assessed. Birth weight was analyzed using birth weight for gestational age z-scores continuously or classified as small or large size for gestational age, but not in analyses that excluded gestational age adjustment. Among girls but not boys, compared to concentrations below detection, a high level of detectable BCEP was associated with significantly shorter mean gestational age and increased odds of preterm birth. In the population overall, a low concentration of detectable BCEP was associated with significantly reduced odds of late/post term birth; the association with high concentration was null. In addition, overall, a high concentration of detectable BCEP was associated with significantly lower odds of small for gestational age (SGA) birth. Both low and high detectable BCEP were associated with reduced odds of SGA among boys. Among girls, low-level detectable BCEP was associated with increased odds of large for gestational age (LGA) birth. Findings from this study suggest that effects of BCEP may be sex-specific and that BCEP may reduce the length of gestation particularly in girls but may not reduce birth weight relative to the duration of gestation.		

Overall Quality Determination **Medium**

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants in this study were all recruited from cohorts that were enrolled in the "Environmental influences on Child Health Outcomes" (ECHO) program at the NIH, where existing cohorts enrolled their participants into the ECHO program set of cohorts. All cohorts included were birth cohorts with urine samples from pregnant mothers and measures of birth outcomes. Individual cohorts followed their own protocols prior to 2019, but in late 2019 ECHO instituted a common protocol for cohorts to follow. Data for this study was collected prior to initiation of the common protocol. In total 16 separate cohorts contributed data to this study, with data collected and submitted to ECHO prior to March 2022. Study details are available for each cohort, including the number of participants included, a description of the sample, eligibility criteria, and their strategy for obtaining birth outcome information. Of the ECHO cohort participants eligible based on the availability of maternal urine samples for organophosphate ester (OPE) quantification (n=12,873), n=7,048 (54.8%) participants had urinary OPE and dilution data. The number of urine samples included and excluded from each cohort was provided. Participants were excluded if they had no information on children (n=82), had multiple births (n=10), or had missing outcome data (n=309). One additional child born with a gestational age of >42 weeks was also excluded, as the formula used for calculating birth weight for gestational age z-scores did not accommodate gestations of that duration. In total, the final sample was n=6,646 mother-child dyads. The study compared the characteristics of participants included to those excluded; while no statistical analysis was done across these two groups, there do not appear to be significant differences in characteristics (e.g. socio-demographics, BMI, tobacco use) of included vs. excluded participants. Overall, key study information for the overall ECHO program and for individual cohorts was well-reported and indicated that significant selection bias was unlikely.
	Metric 2:	Attrition	Medium	The study indicates that 309 out of 7,048 mother-child dyads with urinary OPE measures had missing birth outcome data (gestational age, birth weight, biological sex at birth) and were thus excluded from the study. Another 82 pregnancies had no available child information, and 11 additional children were ineligible (multiple births, length of gestation). The study does not provide details as to why some dyads were missing outcome data, but there is no evidence that any cause would be related to their outcome status.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 3:	Comparison Group	Medium	Since this study was a pooled analyses of multiple birth cohorts, participants were inherently drawn from different background populations and locations across the United States. Key demographic and study design characteristics that may have differed across cohorts were considered as potential covariates (e.g., maternal race/ethnicity, education, age, marital status, along with timing, season, and year of sample collection). Clustering by cohort was addressed in the statistical model. The authors did not discuss adjusting for region or examining potential modification by cohort or region. Instead, to examine the potential influence of heterogeneity across cohorts, the authors included a sensitivity analysis using a "leave-one-cohort-out" approach to assess the influence of each cohort. While there was generally little or no impact of leaving out any one cohort, for the TCEP metabolite BCEP (alternative acronym BCETP) omitting one of the two largest cohorts (CANDLE, predominantly African American mothers from Memphis, TN) strengthened a positive association between low concentrations and birth weight z-scores, which became statistically significant. While there was a change in magnitude and statistical significance, there was no change in direction, and no evidence of bias due to cohort heterogeneity.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	In all participants, TCEP was measured via the urinary metabolite BCEP. Urine samples were collected between 2007 and 2020. Measures used a single spot or first morning urine samples collected from each participant primarily during the second (44.1%) and third (55.5%) trimesters of pregnancy. BCEP is an appropriate biomarker of exposure to TCEP. BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control measures are well-detailed. The LOD was specified to be 0.02 ng/mL and the percentage of samples above the LOD was specified to be 69%. Machine readings were used for values below LOD. Urinary dilution was addressed using creatinine or specific gravity, depending on the cohort. Exposure measures at only a single time point may not reflect the full scope of TCEP exposure across pregnancy due to variability. In addition, other studies examining TCEP biomarkers reported detecting unmetabolized TCEP in 37% of samples (e.g., Hou et al. 2020 HEROID 10143372); it is uncertain whether or to what extent excluding this factor may introduce error. Though some error is likely, there was no evidence of important error or bias in exposure estimates.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 5:	Exposure Levels	Medium	While the 25th percentile of exposure was below the limit of detection, the median value was 0.52 ng/mL and the 75th and 95th percentiles were 1.58 and 8.22 ng/mL, respectively, which should allow for sufficient contrast compared to participants with no detectable BCEP exposure. In statistical analyses, BCEP was categorized into three groups: undetected, and then low or high detectable exposure depending on whether measurements fell above or below the median of dilution-adjusted samples above the LOD.
	Metric 6:	Temporality	Medium	Temporality is established as exposure was measured during the second or third trimester of pregnancy for outcomes that were measured at birth. However, it is uncertain whether exposures during the second or third trimester reflect the etiologically relevant time window for birth outcome effects in relation to TCEP exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes collected prior to 2019 were ascertained using cohort-specific protocols; the majority of cohorts used medical record abstraction for outcomes, with a small proportion of parent reporting. Children born in or after 2019 had outcomes ascertained following the universal ECHO protocol of assigning outcomes based on medical record abstraction.All instances of birth weight were measured using medical records. Gestational age was largely abstracted from or estimated using data from medical records; methods used by hospitals to estimate gestational age were not always specified. Some cohorts specified using 1st or 2nd trimester ultrasound to calculate gestational age, using last menstrual period in cases of missing data. While there was some variation in the estimation of gestational age, there was no evidence that this variation might significantly impact effect estimates. Gestational age was further categorized as preterm (<37 weeks), early term (37-38 weeks), full term (39-40 weeks), and late/post-term (41-42 weeks). Sex-specific birth weight for gestational age z-scores were calculated based on a United states reference population (Aris et al. 2019, not in HERO; PMID: 31201230). Birth weight for gestational age was also categorized as small for gestational age (<10th percentile) and large for gestational age (>90th percentile), as well as term low birth weight (birth weight of <2,500 g at >=37 weeks).
	Metric 8:	Reporting Bias	Medium	All specified outcomes and results that were outlined in the methods were reported in the results. Effect estimates are presented with confidence intervals. Tabulations of outcomes by level of exposure were not provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Potential covariates were selected using a DAG to identify ad include confounders and precision variables, and to exclude intermediates. Covariates included were maternal race/ethnicity, maternal age at delivery, maternal education, maternal marital status, maternal pre-pregnancy BMI, maternal smoking during pregnancy, parity, infant sex, time of urine sample, season of urine sample, and year of urine sample. A potential limitation is that the authors did not mention evaluating the influence of gestational age adjustment on associations with size at birth. Though gestational age might be an intermediate or collider, there was no direct evidence of important error or bias.
	Metric 10:	Covariate Characterization	Medium	Covariates were collected separately by each cohort within the ECHO program and then harmonized by the ECHO Data Analysis Center. While it is not specified how each of these covariates were measured, there is no evidence that insensitive instruments/methods/measures were used.
	Metric 11:	Co-exposure Counfounding	Medium	The study considered co-exposure to other organophosphate esters, including TPHP, TBUP, TIBP, TDCPP, TBOEP, TCPP, TMPP, TEHP, and TPRP. These co-exposures were measured via urinary metabolites at the same time and through the same process as BCEP. Spearman correlations between BECP and these biomarkers were <0.20. These co-exposures were considered in sensitivity analyses by simultaneously adjusting for these OPEs, and results were similar.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between prenatal concentrations of TCEP metabolites in urine and birth outcomes. The use of linear and logistic regression for continuous and categorical outcomes respectively is appropriate to address the research question. Regressions were performed using generalized estimating equations accounting for clustering at the cohort level, which is appropriate given the fact that cohorts were combined into one group in this study. Potential non-linear associations were examined by categorizing BCEP into three groups of similar size.
	Metric 13:	Statistical Power	Medium	The number of participants in the study (n=6,646 mother-child dyads) is likely sufficiently large to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric		Rating	Comments
	Metric 15:	Statistical Analysis	High	Risk estimates were calculated using linear regression models for continuous outcomes and multinomial logistic regression models for categorical outcomes. BCEP concentrations were categorized, with values below LOD as the referent, and a lower and higher exposure group defined by categorizing remaining samples at their median. Inclusion of covariates was dependent on a DAG created to exclude variables that may be causal intermediates. Covariates with <20% of missing values were addressed via multiple imputation by chained equations. Sensitivity analyses are detailed, including sex stratification, a leave-one-cohort-out approach, and a co-pollutant model.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Urinary concentrations of BCEP are a valid and appropriate biomarker of exposure to TCEP.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 0.02 ng/mL. BCEP was detected in 69% of samples. There was no evidence that the detection limit was inadequate.
	Metric 19:	Biomarker Stability	Medium	The study reports that urine samples were shipped on dry ice to the Human Health Exposure Analysis Resource (HHEAR) at the NYU Grossman School of Medicine. While no other specific storage information is provided, there is no evidence that inappropriate methods were used and the biomarker is not expected to have a high likelihood of stability.
	Metric 20:	Sample Contamination	Medium	No information is provided potential contamination.
	Metric 21:	Method Requirements	Medium	BCEP was quantified using high-performance liquid chromatography with triple quadrupole mass spectrometry, and quality control (QC) measures are well-detailed. One QC measure included the use of blinded duplicate samples (n=127 samples for BCEP). For BCEP, 66% of duplicate pairs were both above LOD; the remaining pairs were not concordant with respect to detection. Details on the magnitude of disparities in estimated concentrations were not provided.
	Metric 22:	Matrix Adjustment	Medium	BCEP was adjusted for urinary dilution, measured as either specific gravity or creatinine depending on the source cohort, using published methods. Results were presented only for dilution-standardized exposure variables.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Oh, J., Buckley, J. P., Li, X., Gachigi, K. K., Kannan, K., Lyu, W. J., Ames, J. L., Barrett, E. S., Bastain, T. M., Breton, C. V., Buss, C., Croen, L. A., Dunlop, A. L., Ferrara, A., Ghassabian, A., Herbstman, J. B., Hernandez-Castro, I., Hertz-Picciotto, I., Kahn, L. G., Karagas, M. R., Kuiper, J. R., McEvoy, C. T., Meeker, J. D., Morello-Frosch, R., Padula, A. M., Romano, M. E., Sathyanarayana, S., Schantz, S., Schmidt, R. J., Simhan, H., Starling, A. P., Tylavsky, F. A., Volk, H. E., Woodruff, T. J., Zhu, Y. Y., Bennett, D. H. (2024). Associations of organophosphate ester flame retardant exposures during pregnancy with gestational duration and fetal growth: The Environmental Influences on Child Health Outcomes (ECHO) program. Environmental Health Perspectives 132(1):017004.
Health Outcome(s) and Reported Health Effect(s):	Reproductive /Developmental-Birth anthropometry (birth weight for gestational age z-score, term low birth weight, small for gestational age, large for gestational age)Length of gestation (gestational age at birth, preterm birth, early term birth, late/post-term birth)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis(2-chloroethyl) phosphate (BCETP)
Linked HERO ID(s):	No linked references.
HERO ID:	11577123

Domain	Metric	Rating	Comments
Additional Comments:	This prospective study used data from 16 cohorts participating in the ECHO (Environmental influences on Child Health Outcomes) program to estimate the association between maternal urinary organophosphate ester concentrations and two types of birth outcome: birth weight and duration of gestation. Spot urine concentrations of BCEP from the second or third trimester were used to estimate exposure. There were no major concerns for bias. There was some degree of uncertainty due to the use of only a single timepoint measurement of exposure and variability in ways in which gestational age was assessed. Birth weight was analyzed using birth weight for gestational age z-scores continuously or classified as small or large size for gestational age, but not in analyses that excluded gestational age adjustment. Among girls but not boys, compared to concentrations below detection, a high level of detectable BCEP was associated with significantly shorter mean gestational age and increased odds of preterm birth. In the population overall, a low concentration of detectable BCEP was associated with significantly reduced odds of late/post term birth; the association with high concentration was null. In addition, overall, a high concentration of detectable BCEP was associated with significantly lower odds of small for gestational age (SGA) birth. Both low and high detectable BCEP were associated with reduced odds of SGA among boys. Among girls, low-level detectable BCEP was associated with increased odds of large for gestational age (LGA) birth. Findings from this study suggest that effects of BCEP may be sex-specific and that BCEP may reduce the length of gestation particularly in girls but may not reduce birth weight relative to the duration of gestation.		

Overall Quality Determination **Medium**

Study Citation:	Percy, Z., Chen, A., Sucharew, H., Yang, W., Vuong, A. M., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Cecil, K. M., Xu, Y., Yolton, K. (2023). Early-life exposure to a mixture of organophosphate esters and child behavior. International Journal of Hygiene and Environmental Health 250:114162.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Behavioral Assessment System for Children, 2nd Edition (BASC-2): scales for internalizing problems, externalizing problems, and Behavioral Symptom Index (BSI)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581667

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	This cohort study examined organophosphate ester (OPE) mixtures and child behavior among participants recruited between March 2003 and February 2006 for the Health Outcomes and Measures of the Environment (HOME) Study, a pregnancy and birth cohort based in Cincinnati, Ohio. Eligibility criteria for pregnant women included: being at least 18 years of age, being at 16 +/- 3 weeks of gestation, living in a home built before 1978, being fluent in English, not being diagnosed with various conditions (bipolar disorder, schizophrenia, diabetes, or cancer), not taking medications for thyroid disorder of seizures, being HIV negative, and not planning to move outside of the Greater Cincinatti Area. Participants provided urine samples that were used to estimate organophosphate ester (OPE) metabolites at multiple study visits during pregnancy and early childhood. For this study of organophosphate exposure and child behavior, the authors included mother-child pairs with at least one measure of urinary OPE metabolites from pregnancy or childhood, as well as at least one parent-completed behavioral assessment at age 3 or 8 years. Of 389 women who gave birth to singletons, 170 did not meet these criteria, resulting in a final analytical sample of 219 mother-child pairs. Included and excluded participants were not compared, but there was no evidence to suggest inclusion might be associated with OPE exposure. The authors provided sufficient information pertaining to participant selection, and there are no major concerns of selection bias.
	Metric 2:	Attrition	Medium	Of 389 pregnant women in the HOME cohort who gave birth to singleton infants, 170 were excluded because they did not have at least one measure of OPEs for the mother or child, or at least one parent assessment of children using the Behavioral Assessment System for Children (BASC-2) at age 3 or 8 years. There were no additional exclusions due to missing covariate data. The analysis sample of 219 children included 56.2% of singleton births. While there was attrition, there was no evidence of differences in included vs. excluded participants that might induce bias. No major concerns were noted.
	Metric 3:	Comparison Group	Medium	All participants were recruited from the same eligible population using the same criteria. The authors provided clear details about the study setting and key inclusion and exclusion details. There was no evidence for concerns related to the adequacy of the comparison group of participants with lower levels of OPE metabolites.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Sucharew, H., Yang, W., Vuong, A. M., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Cecil, K. M., Xu, Y., Yolton, K. (2023). Early-life exposure to a mixture of organophosphate esters and child behavior. International Journal of Hygiene and Environmental Health 250:114162.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Behavioral Assessment System for Children, 2nd Edition (BASC-2): scales for internalizing problems, externalizing problems, and Behavioral Symptom Index (BSI)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581667

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Spot urines or urine samples collected in diaper inserts were used to measure OPE metabolite concentrations that included the TCEP metabolite BCEP. While associations were analyzed for an OPE mixture and not for TCEP individually, there were no notable limitations with respect to the measurement approach. Samples were collected from participating mothers (weeks 16 and 26 of gestation, at delivery) and children (ages 1, 2, 3 and 5 years). The number of samples collected varied by participant. While the number of samples per mother-child pair was not reported, sample sizes across time points ranged from n=153 to n=217. Intra-class correlations (ICC) for repeated measures were low (ICC = 0.19 for children, 0.27 for mothers), indicating that there was substantial variation in concentrations within individuals over time. BCEP was quantified using HPLC with mass spectrometry. The authors reported an LOD of 0.1 ug/L, and the percent of samples with detectable levels ranged from 83.9% at 26 weeks of gestation to 95.5% in children at age 1 year. Values below LOD were imputed as LOD divided by the square root of 2. To account for dilution, laboratory technicians measured the specific gravity of all urine samples using an ATAGO PAL-10S pocket refractometer. Although exposure at each time point was represented by a single sample, an important strength of this study was the availability of repeated measures from both pregnancy and early childhood.
	Metric 5:	Exposure Levels	Medium	Table 1 provided medians and IQRs for individual OPE metabolites, including BCEP, at each timepoint. The median (IQR) in ug/L at was 0.60 (0.34, 1.07) at 16 weeks' gestation, and 1.13 (0.47, 2.73) at age 1 year. The range and distribution of exposure reported in this study appeared sufficient to develop an exposure-response estimate, and the authors utilized a continuous measure of exposure for their analyses. However, it was a limitation that associations between OPE exposures and behavioral outcomes were presented only for metabolite mixtures, and not for individual metabolites.
	Metric 6:	Temporality	High	Temporal sequencing was appropriate. OPE exposure measures from pregnancy were used to estimate associations with behavioral outcomes measured at both ages 3 and 8 years. Exposure measures from childhood (ages 1 to 5 years) were used to estimate associations with outcomes at age 8 years. While the importance of specific exposure windows for BCEP and child neurodevelopment is unclear, analysis of both prenatal and postnatal exposure measures, along with appropriate sequencing, were strengths of this study.

Domain 3: Outcome Assessment

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Sucharew, H., Yang, W., Vuong, A. M., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Cecil, K. M., Xu, Y., Yolton, K. (2023). Early-life exposure to a mixture of organophosphate esters and child behavior. International Journal of Hygiene and Environmental Health 250:114162.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Behavioral Assessment System for Children, 2nd Edition (BASC-2): scales for internalizing problems, externalizing problems, and Behavioral Symptom Index (BSI)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581667

Domain	Metric		Rating	Comments
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes were assessed using the BASC-2 questionnaire, which was completed by caregivers (98% by mothers) at the 3- or 8-year study visits after receiving instructions from study staff. The BASC-2 is widely used 160-item parent-reported assessment of children's behavior in public and home settings. This assessment provides four composite scales; three were assessed in this study (Internalizing Problems, Externalizing Problems, and the Behavioral Symptom Index (BSI)). BASC-2 scores are age normalized to a mean of 50 and a SD of 10. The authors cited literature reporting that the internal consistency of the BASC-2 ranges from r = 0.87-0.95 across ages 3 and 8 years, and correlations with other measures of child behavior from r = 0.65-0.84. A potential limitation is that the behavioral measures used in this study did not include assessments by teachers or personnel with specialized training. Some measurement error in parental reporting is likely, but there was no evidence of systematic bias, and models adjusted for variables that included maternal depression and income. There was no evidence of inadequate validity or important bias in outcome assessment.
	Metric 8:	Reporting Bias	High	The authors provided details about the measured outcomes, and associated effect estimates are reported along with 95% confidence intervals. No concerns of reporting bias are noted in this study.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	The authors detailed that a number of variables were considered as potential confounders including breastfeeding status, maternal race, household income at baseline, maternal depression at the time of BASC-2 measurement, maternal education at baseline, maternal age at baseline, marital status at baseline, and the caregiving environment. The final covariate set was selected using a directed acyclic graph, and included maternal depression, breastfeeding status, maternal race (white or non-white), household income, and caregiving environment. The authors did not explicitly state that BASC-2 scores used were normed for child sex as well as child age at administration, and there was no stratification by or adjustment for child sex. Covariates did not include indicators of schooling or behavioral interventions. Nonetheless, there was no direct evidence of important residual confounding or that potential intermediates were included as covariates.
	Metric 10:	Covariate Characterization	Medium	All variables considered as covariates were assessed using questionnaires and methods that included well-established instruments. The caregiving environment was assessed using the Home Observation and Measurement of the Environment, which is a questionnaire and interview obtaining information on the quality and quantity of caregiving in the home at 12 months of age. Maternal depression was measured using the Beck's Depression Inventory. There was no evidence that the methods employed to measure covariates had poor validity.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Sucharew, H., Yang, W., Vuong, A. M., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Cecil, K. M., Xu, Y., Yolton, K. (2023). Early-life exposure to a mixture of organophosphate esters and child behavior. International Journal of Hygiene and Environmental Health 250:114162.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Behavioral Assessment System for Children, 2nd Edition (BASC-2): scales for internalizing problems, externalizing problems, and Behavioral Symptom Index (BSI)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581667

Domain	Metric		Rating	Comments
	Metric 11:	Co-exposure Counfounding	Medium	Exposure was assessed as a mixture of three OPE metabolites: BCEP, BDCIPP, bis(1,3-dichloro-2-propyl) phosphate [BDCIPP], and diphenyl phosphate [DPHP]. However, correlations among these OPE metabolites were not reported, and potential confounding of associations with one metabolite by others was not directly assessed. The authors cited a study in which an OPE metabolite not measured in this study – isopropyl-phenyl phenyl phosphate (ip-PPP) – was associated with child behavior at age 8 years. The authors noted the possibility that confounding by co-exposures as a potential concern. However, there was no direct evidence of such bias.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design implemented was appropriate for the research question being examined. Urinary OPE concentrations were measured prenatally and in early life, and associations with BASC-2 scores in children at later ages were examined. The authors also applied appropriate statistical methods to evaluate the primary research questions. No concerns were noted for the general study design.
	Metric 13:	Statistical Power	Medium	There was variability in both exposure measures; statistical power was increased by analyzing both exposure and outcome using continuous variables, as well as the availability of repeated measures. There was no evidence that the number of participants in this study (n=219 mother-child pairs) was inadequate to detect associations.
	Metric 14:	Reproducibility of Analyses	Medium	The authors provided a sufficient description of their analytical methods to allow others to conceptually reproduce their results with access to the pertinent data. No major concerns are noted, as the description of analyses performed was clear.
	Metric 15:	Statistical Analysis	Low	Descriptive data were provided for participant characteristics and OPE exposure variables, but not for BASC-2 scores. In this paper, the authors estimated associations between mixtures of natural log transformed OPE metabolite concentrations and behavioral outcomes using two methods: structural equation modeling and quantile g-computation. These analysis methods are sound for the analysis of mixtures. Analyses of associations between behavioral outcomes and individual OPEs were not discussed. The exclusive analysis of mixtures negates the utility of the study for evaluating potential effects specific to TCEP or other individual OPEs. Factor loadings for the latent OPE exposure variable provided limited information suggesting that BCEP and BDCIPP were potential drivers of associations. BCEP was the reference chemical against which factor loadings for BDCIPP and DPHP were reported; loadings were highest for BCEP at age 2 years, and lower for BCEP than for BDCIPP at other time points. However, associations for OPE mixtures at age 2 years were not significant; patterns of associations were inconsistent over time. While there were no major concerns for the analysis of OPE mixtures, the analysis strategy provided little meaningful information on potential effects of TCEP individually, warranting a low rating.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Sucharew, H., Yang, W., Vuong, A. M., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Cecil, K. M., Xu, Y., Yolton, K. (2023). Early-life exposure to a mixture of organophosphate esters and child behavior. International Journal of Hygiene and Environmental Health 250:114162.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Behavioral Assessment System for Children, 2nd Edition (BASC-2): scales for internalizing problems, externalizing problems, and Behavioral Symptom Index (BSI)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581667

Domain	Metric		Rating	Comments
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Concentrations of the TCEP metabolite BCEP were quantified in urine samples obtained from mothers or their children. Multiple samples were obtained from participants, though the mean and range in number of repeated measures was not provided in this manuscript. A single urine sample was obtained for each period of interest; ICCs for repeated measures were 0.27 for prenatal and 0.19 for postnatal BCEP measures.
	Metric 17:	Effect Biomarker	N/A	No biomarkers of effect were assessed.
	Metric 18:	Method Sensitivity	Medium	The authors reported the LOD for TCEP as 0.1 ug/L, and this LOD was low enough to detect TCEP in a sufficient percentage of samples. The authors reported the percent of samples with detectable BCEP as ranging from 83.9% to 95.5%. No concerns of method sensitivity are noted.
	Metric 19:	Biomarker Stability	Medium	The authors note that samples were aliquoted and stored at -20 degrees Celsius until they were shipped overnight to the CDC laboratory. Further details were not provided. No concerns of biomarker stability were noted.
	Metric 20:	Sample Contamination	Medium	The authors did not discuss the potential for contamination in this study. However, parent compounds were not measured, and there was no evidence of inappropriate handling or processing that might lead to contamination of samples with OPE metabolites.
	Metric 21:	Method Requirements	Medium	Appropriate methods were implemented to quantify exposure concentrations in urine samples including automated off-line solid-phase extraction, reversed phase high-performance liquid chromatography, and isotope dilution-electrospray ionization tandem mass spectrometry. Total free and conjugated concentrations of metabolites were measured. A reference publication with further analytic details and quality control methods were cited. No concerns are noted pertaining to the exposure quantification methods.
	Metric 22:	Matrix Adjustment	Medium	The authors reported their results using specific gravity adjusted urinary concentrations of TCEP. All results reported are associated with this matrix-adjusted concentration.

Additional Comments: This study examined participants from the Health Outcomes and Measures of the Environment (HOME) Study, a pregnancy and birth cohort from Cincinnati, Ohio. It examined the association between urinary concentrations of a mixture of OPE metabolites that included BCEP, a metabolite of TCEP, and scores from several components of BASC-2. The final analysis examined 219 mother-child pairs. A strength of this study is that authors obtained urine samples for OPE exposure quantification at several timepoints. Associations with a mixture of three OPE metabolites measured at multiple timepoints in pregnancy and in early childhood were reported. The direction and magnitude of associations varied, and few were statistically significant. The authors provided sufficient details pertaining to numerous aspects of their methodology. While this study had important strengths, the lack of analyses of associations with individual OPE metabolites was a major concern. All associations were reported for a mixture of three OPE metabolites, making it impossible to determine the extent to which there were any TCEP-specific effects. The study did not provide associations with TCEP for data extraction.

Overall Quality Determination Medium

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	All key elements of the study design were present. Risk of participant and selection bias is low; methods of participant selection and inclusion/exclusion criteria are well-stated. The exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study.
	Metric 2:	Attrition	Medium	Moderate exclusion from the analysis sample. Excluded participant demographics (i.e., missing IQ data) were presented and exclusion reasoning was adequately addressed.
	Metric 3:	Comparison Group	High	Key elements of the study design are reported (i.e., setting, inclusion/exclusion criteria, and methods of participant selection), and subjects were indicated as similar by population and timeframe recruitment using the same inclusion/exclusion criteria and were of similar age and health status. Potential confounding and stratification variables were addressed and controlled by statistical analysis including effect modification by both individual- and neighborhood-level SES.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Exposure was consistently assessed via spot urine and blood sample using well-established methods, sampling strategy was adequately described, and LOD was stated. Slight potential for exposure misclassification given that method of urine collection was based on participant's toilet-training status, but not expected to change effect estimate.
	Metric 5:	Exposure Levels	Medium	The range and distribution of exposure is sufficient to develop an exposure-response estimate and 3 or more levels of exposure (i.e., ages 1, 2, 3, and 5) are reported.
	Metric 6:	Temporality	Medium	Temporality for exposure is established (maternal and early life exposure), but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The outcome was assessed using well-established methods stated as "gold standard". Researchers assessing participant cognitive abilities were blinded, trained, and evaluated without notice every six months. Low concern for reporting biases or outcome misclassification.
	Metric 8:	Reporting Bias	High	A description of all measured outcomes are reported in the methods and abstract that allows for detailed extraction. Effect estimates are reported with a 95% confidence interval, and sample size information to be used in exposure-response or descriptive analyses are presented.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Specific covariates were discussed and adjustments were made for potential confounders (e.g. age, sex, socioeconomic status) in the final analyses through the use of statistical models including adjustment in multivariate models and stratification. All methods to address confounding are well-documented.
	Metric 10:	Covariate Characterization	High	Potential confounders were assessed using valid and reliable methodology where appropriate. SES was assessed at both the individual-level and neighborhood-level. Assessment methodology and reasoning was well-documented.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposures to pollutant that are not the target exposure were not addressed in the study.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design chosen was appropriate to address the research question to assess the associations between exposure to three OPE metabolites (including BCEP) and cognitive ability in children. Statistical methods appropriate to address the research question were used, including repeated measures analyses and regression modeling.
	Metric 13:	Statistical Power	Medium	The number of participants are adequate to detect an effect in the exposed population.
	Metric 14:	Reproducibility of Analyses	Medium	Reproducibility of analyses is likely given description of analysis.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described, and the data fit normality after adjustments.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP biomarker of interest is metabolite of TCEP.
	Metric 17:	Effect Biomarker	N/A	Biomarkers of effect were not assessed.
	Metric 18:	Method Sensitivity	Low	LOD is sufficient and analytical methods for measuring biomarker are adequately reported. LOQ is not reported.
	Metric 19:	Biomarker Stability	Medium	Samples appeared to have been handled correctly during collection, processing and storage.
	Metric 20:	Sample Contamination	Medium	Documentation of sampling is provided, but no information on contamination is reported.
	Metric 21:	Method Requirements	Medium	Instrumentation that allows for identification of the biomarker with a high degree of confidence and the required sensitivity were used.
	Metric 22:	Matrix Adjustment	Medium	Only adjusted results are provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric	Rating	Comments
Additional Comments:	This study did not find any statistically significant associations between early-life urinary OPE metabolites and cognitive abilities at 8 years old. This study collected urine from participants of a larger birth cohort study (HOMES) at ages 1, 2, 3, and 5. Blood was collected at age 5 and cognitive abilities were measured at age 8 through the administration of an IQ test. Maternal IQ was also tested at the same time as the child participant. Biomonitoring media and IQ collections and analytical processes were well documented and seemingly valid. Various potential covariates and confounding variables were considered and adjusted for.		

Overall Quality Determination High

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	All key elements of the study design were present. Risk of participant and selection bias is low; methods of participant selection and inclusion/exclusion criteria are well-stated. The exposure-outcome distribution of the participants is likely representative of the exposure-outcome distributions in the population of persons eligible for inclusion in the study.
	Metric 2:	Attrition	Medium	Moderate exclusion from the analysis sample. Excluded participant demographics (i.e., missing IQ data) were presented and exclusion reasoning was adequately addressed.
	Metric 3:	Comparison Group	High	Key elements of the study design are reported (i.e., setting, inclusion/exclusion criteria, and methods of participant selection), and subjects were indicated as similar by population and timeframe recruitment using the same inclusion/exclusion criteria and were of similar age and health status. Potential confounding and stratification variables were addressed and controlled by statistical analysis including effect modification by both individual- and neighborhood-level SES.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Exposure was consistently assessed via spot urine and blood sample using well-established methods, sampling strategy was adequately described, and LOD was stated. Slight potential for exposure misclassification given that method of urine collection was based on participant's toilet-training status, but not expected to change effect estimate.
	Metric 5:	Exposure Levels	Medium	The range and distribution of exposure is sufficient to develop an exposure-response estimate and 3 or more levels of exposure (i.e., ages 1, 2, 3, and 5) are reported.
	Metric 6:	Temporality	Medium	Temporality for exposure is established (maternal and early life exposure), but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The outcome was assessed using well-established methods stated as "gold standard". Researchers assessing participant cognitive abilities were blinded, trained, and evaluated without notice every six months. Low concern for reporting biases or outcome misclassification.
	Metric 8:	Reporting Bias	High	A description of all measured outcomes are reported in the methods and abstract that allows for detailed extraction. Effect estimates are reported with a 95% confidence interval, and sample size information to be used in exposure-response or descriptive analyses are presented.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Specific covariates were discussed and adjustments were made for potential confounders (e.g. age, sex, socioeconomic status) in the final analyses through the use of statistical models including adjustment in multivariate models and stratification. All methods to address confounding are well-documented.
	Metric 10:	Covariate Characterization	High	Potential confounders were assessed using valid and reliable methodology where appropriate. SES was assessed at both the individual-level and neighborhood-level. Assessment methodology and reasoning was well-documented.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposures to pollutant that are not the target exposure were not addressed in the study.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design chosen was appropriate to address the research question to assess the associations between exposure to three OPE metabolites (including BCEP) and cognitive ability in children. Statistical methods appropriate to address the research question were used, including repeated measures analyses and regression modeling.
	Metric 13:	Statistical Power	Medium	The number of participants are adequate to detect an effect in the exposed population.
	Metric 14:	Reproducibility of Analyses	Medium	Reproducibility of analyses is likely given description of analysis.
	Metric 15:	Statistical Analysis	High	Model assumptions were adequately described, and the data fit normality after adjustments.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	BCEP biomarker of interest is metabolite of TCEP.
	Metric 17:	Effect Biomarker	N/A	Biomarkers of effect were not assessed.
	Metric 18:	Method Sensitivity	Low	LOD is sufficient and analytical methods for measuring biomarker are adequately reported. LOQ is not reported.
	Metric 19:	Biomarker Stability	Medium	Samples appeared to have been handled correctly during collection, processing and storage.
	Metric 20:	Sample Contamination	Medium	Documentation of sampling is provided, but no information on contamination is reported.
	Metric 21:	Method Requirements	Medium	Instrumentation that allows for identification of the biomarker with a high degree of confidence and the required sensitivity were used.
	Metric 22:	Matrix Adjustment	Medium	Only adjusted results are provided.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Percy, Z., Chen, A., Yang, W., Braun, J. M., Lanphear, B., Ospina, M., Calafat, A. M., Xie, C., Cecil, K. M., Vuong, A. M., Xu, Y., Yolton, K. (2022). Childhood urinary organophosphate esters and cognitive abilities in a longitudinal cohort study. Environmental Research 215(Pt 1):114265.
Health Outcome(s) and Reported Health Effect(s):	Neurological/Behavioral-Cognitive abilities/IQ (WISC, FSIQ)
Chemical:	Tris(2-chloroethyl) phosphate (TCEP)- Metabolite: Bis-2-chloroethyl phosphate (BCEP)
Linked HERO ID(s):	No linked references.
HERO ID:	11581664

Domain	Metric	Rating	Comments
Additional Comments:	This study did not find any statistically significant associations between early-life urinary OPE metabolites and cognitive abilities at 8 years old. This study collected urine from participants of a larger birth cohort study (HOMES) at ages 1, 2, 3, and 5. Blood was collected at age 5 and cognitive abilities were measured at age 8 through the administration of an IQ test. Maternal IQ was also tested at the same time as the child participant. Biomonitoring media and IQ collections and analytical processes were well documented and seemingly valid. Various potential covariates and confounding variables were considered and adjusted for.		

Overall Quality Determination High