

Protocol

Investigating the Association Between Volatile Organic Compounds and Preterm Birth in Detroit, Michigan: Protocol for the Center for Leadership in Environmental Awareness and Research Birth Cohort Study

Andrea E Cassidy-Bushrow^{1,2}, PhD; Emily Leydet¹, MPH; Albert M Levin^{1,3}, PhD; Mei Lu^{1,3}, PhD; Jennifer K Straughen^{4,5}, PhD

¹Department of Public Health Sciences, Henry Ford Hospital, Detroit, MI, United States

²Department of Pediatrics and Human Development, College of Human Medicine, Michigan State University, East Lansing, MI, United States

³Department of Epidemiology and Biostatistics, College of Human Medicine, Michigan State University, East Lansing, MI, United States

⁴Department of Women's Health Services, Henry Ford Hospital, Detroit, MI, United States

⁵Department of Obstetrics, Gynecology and Reproductive Biology, College of Human Medicine, Michigan State University, East Lansing, MI, United States

Corresponding Author:

Andrea E Cassidy-Bushrow, PhD
Department of Public Health Sciences
Henry Ford Hospital
1 Ford Place
Detroit, MI 48202
United States
Phone: 1 3138746097
Email: acassid1@hfhs.org

Abstract

Background: Preterm birth (PTB), or birth before 37 weeks of gestation, remains a significant public health issue in the United States, particularly in Detroit, Michigan. Growing evidence suggests that volatile organic compounds (VOCs), aromatic or chlorinated organic compounds that vaporize readily, may influence PTB risk. However, much of this prior work is limited by indirect VOC exposure estimates (eg, assignment based on maternal residential address), single-point or cumulative exposure estimates during pregnancy, or limited consideration of potential mechanistic factors.

Objective: The Center for Leadership in Environmental Awareness and Research (CLEAR) birth cohort has been designed to test the hypotheses that prenatal VOC exposures, measured as VOC metabolites in maternal urine, increase the risk of PTB; that VOC exposures are associated with maternal inflammation and placental function measures; that associations between prenatal VOC exposures and PTB may be mediated by these maternal inflammation and placental function measures; and that there are neighborhood-level factors that may increase the risk of VOC exposure during pregnancy.

Methods: A prospective cohort of 1075 pregnant patients receiving prenatal care at Henry Ford Health will be recruited. Pregnant patients residing in Detroit or receiving prenatal care at a Detroit-based Henry Ford Health women's health clinic are eligible. Pregnant patients are followed until delivery. Up to 3 urine and blood samples (collected during early, mid, and late pregnancy) are obtained for measurement of VOC metabolites and inflammatory biomarkers, respectively. The placenta is obtained after delivery for epigenomic and transcriptomic measurement. Surveys are administered to pregnant participants to assess a variety of lifestyle, psychosocial, medical, residential, and other factors. Address information collected from both surveys and electronic medical records across pregnancy will be used to identify potential sources of VOC exposure. The electronic medical record is used to obtain medical and delivery data, including infant sex, date of delivery, and gestational age (GA) at delivery. PTB, the primary study outcome, is defined as GA at delivery <37 weeks. A nested case-control approach (frequency matching PTB cases 1:1 with full-term controls [GA at delivery ≥37 weeks] based on infant sex and maternal race) will be applied. Statistical methods, including logistic regression, linear mixed methods, and geographically weighted regression models, as well as chemical mixture approaches, will be used.

Results: Funding began September 2022 and recruitment commenced November 2023. Through April 22, 2026, a total of 468 pregnant patients have consented to participate in the CLEAR birth cohort, and recruitment is ongoing.

Conclusions: The CLEAR cohort will provide novel data on the role of VOCs during pregnancy in the risk of PTB. Additionally, the role of VOC exposures during pregnancy in maternal inflammation and placental function will be examined. Finally, potential sources of VOC exposures, which could be targets for environmental remediation, will be identified.

International Registered Report Identifier (IRRID): DERR1-10.2196/87272

JMIR Res Protoc 2026;15:e87272; doi: [10.2196/87272](https://doi.org/10.2196/87272)

Keywords: birth cohort; volatile organic compounds; preterm birth; placenta; inflammation

Introduction

Preterm birth (PTB; gestational age [GA] at delivery <37 weeks) is associated with both immediate and long-term health consequences for the infant, including increased risks of infant mortality, neurodevelopmental delays, obesity, and other chronic diseases. The United States consistently reports higher PTB rates than other high-income nations, with certain cities (particularly those in the postindustrial “rust belt” along the Great Lakes waterways) exceeding the national average. Among these, Detroit, Michigan, has the highest PTB rate of all large US cities, significantly surpassing the national rate (15.6% vs 10.4%) [1,2]. Despite this, few studies have sufficiently examined the contribution of environmental contaminants such as volatile organic compounds (VOCs)—prevalent in urban areas such as Detroit—to PTB risk. Thus, Detroit, which may be “...paradigmatic of communities around the world that depend on heavy industry” [3], may be an ideal setting for studying such associations [4].

VOCs are a vast group of aromatic or chlorinated organic compounds that vaporize readily at room temperature. Common VOCs include benzene, toluene, ethylbenzene, and xylene (collectively known as BTEX), trichloroethylene, and tetrachloroethylene (also known as perchloroethylene). There is growing evidence demonstrating that BTEX, trichloroethylene, and perchloroethylene increase the risk of PTB [4-13]. In our previous study of Detroit pregnancies, every 5-unit increase in airshed BTEX level was associated with 1.54 times higher odds of PTB (95% CI 1.25-1.89) [14]. Most of these prior studies are limited by reliance on estimated exposures and thus are subject to bias due to potential exposure misclassification.

Beyond enhancing our understanding of the association between VOCs and PTB risk, additional investigations into the mechanisms by which VOCs may lead to PTB are needed, as such studies may inform future prevention efforts. On the basis of a small but growing body of evidence, we hypothesize that VOCs may impact maternal systemic inflammation and DNA methylation. Both benzene and perchloroethylene can alter cytokine levels [15,16]. Similarly, VOC exposure

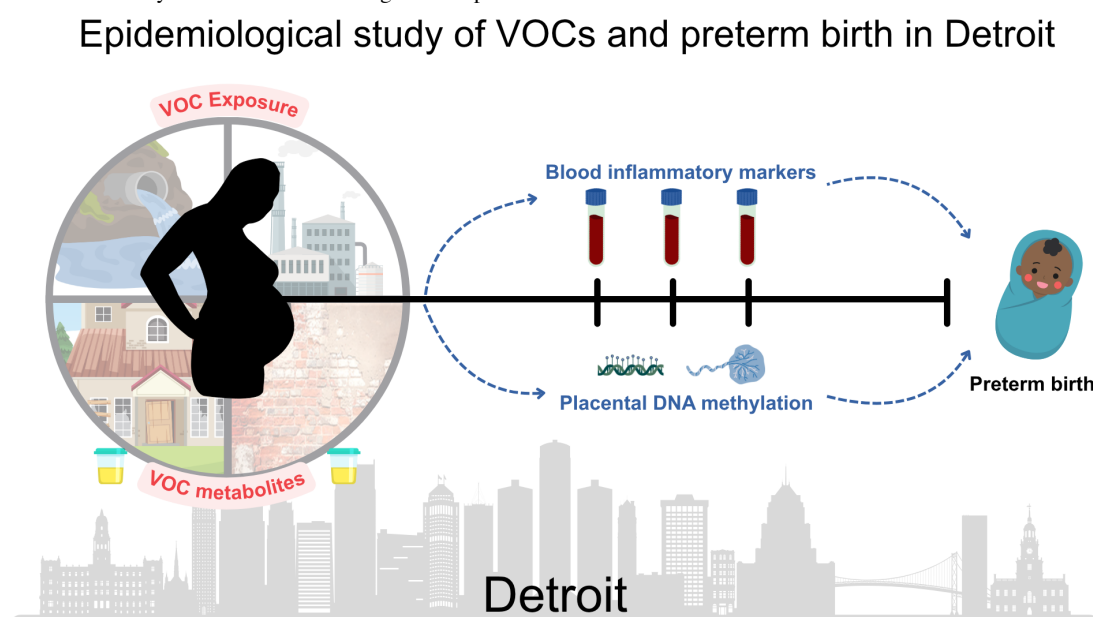
alters DNA methylation patterns and gene expression in nontarget tissue (ie, blood) [17-22]. Our recently published data from Detroit pregnancies corroborate these findings, with higher airshed BTEX levels being associated with higher midpregnancy interleukin-1 beta and tumor necrosis factor-alpha levels [23] and differentially methylated regions [24] in maternal blood. In the Lifestyle-Immune System-Allergy birth cohort, maternal perchloroethylene exposure was associated with fewer interferon-gamma-producing T-cells [25]. Finally, in the EDEN (*Etude des Déterminants pré et post natalis précoces du développement psychomoteur et de la santé de l'ENfant*) birth cohort study, higher maternal benzene exposure was associated with decreased cord blood Cluster of Differentiation 4+ Cluster of Differentiation 25+ T-regulatory cells [26]. Despite these suggestive findings, there is a need for additional studies examining direct VOC exposure (rather than estimates based on residential address or occupation), at multiple time points across pregnancy, and which incorporate a biologically-relevant tissue—the placenta.

To address these gaps, the Center for Leadership in Environmental Awareness and Research (CLEAR) birth cohort study, which is part of a National Institute of Environmental Health Sciences-funded Superfund Research Program P42 Center and focused on 6 specific VOCs (Table 1) and PTB, is underway. The CLEAR P42 Center consists of 2 engineering projects (projects 1 and 2), 3 biomedical projects (projects 3, 4, and 5), and 5 cores (administrative core, chemical analysis core, community engagement core, research experience and training coordination core, and data management and analytics core). The specific aims of the CLEAR birth cohort study are to examine whether VOC metabolite levels in maternal urine are associated with PTB, maternal systemic inflammation, and changes in placental function (as measured by DNA methylation and transcriptional gene expression). Exploratory work considering maternal inflammation and placental function changes as potential mediators of associations between VOCs and PTB is planned. Finally, potential sources of VOC exposures will be examined. An overview of the study aims is shown in Figure 1.

Table 1. Volatile organic compound (VOC) metabolites to be measured in urine samples from participants in the Center for Leadership in Environmental Awareness and Research birth cohort study and major sources of exposure.

Parent VOC	Metabolite	Major sources
Trichloroethylene	N-acetyl-S-(1,2-dichlorovinyl)-L-cysteine and N-acetyl-S-(2,2-dichlorovinyl)-L-cysteine	Extraction solvent; dry cleaning; paints or strippers, adhesives, lubricants, varnish, and pesticides; and cold metal cleaners [27]
Tetrachloroethylene	N-acetyl-S-(trichlorovinyl)-L-cysteine	Dry cleaning industry and liquid waste from metal degreasing [28]
Benzene	S-phenylmercapturic acid and trans,trans-muconic acid	Auto exhaust, tobacco smoke, automobile service stations, and industrial emissions [29]
Toluene	Benzylmercapturic acid	Gasoline or fuel production; paints; paint thinners; fingernail polish, lacquers, adhesives, and rubber; and tobacco [30]
Ethylbenzene	Phenylglyoxylic acid	Styrene production, cleaners, paints, fuel, pesticides, solvents, and tobacco smoke [31, 32]
Xylene	Dimethylphenyl mercapturic acid, 2MHA ^a , and 3MHA+4MHA	Auto exhaust; industrial solvent; airplane fuel; gasoline; paint, varnish, or shellac; rust prevention products; and cigarettes [33]

^aMHA: methylhippuric acid.

Figure 1. Overview of the study aims. VOC: volatile organic compound.

Methods

Participants and Recruitment

CLEAR is a prospective longitudinal birth cohort study that will follow a target sample of 1075 pregnant patients in the Detroit area from pregnancy through delivery. Recruitment for CLEAR began in November 2023 and is ongoing. Through June 2025, participants were enrolled into both the CLEAR birth cohort protocol and a birth cohort focused on asthma and allergy [34-36] under a combined protocol. To reduce participant burden and focus on pregnant patients in Detroit, CLEAR-only enrollment began in July 2025, which reduced data and sample collection requirements. The CLEAR protocol is identical for those enrolled both before and after June 2025.

To be eligible for the CLEAR study, participants must be aged ≥ 18 years, English-speaking, either receiving prenatal care at a Detroit-based Henry Ford Health (HFH) Women's Health Services (WHS) clinic or residing in Detroit and receiving care at any HFH WHS clinic, and planning to

deliver at an HFH hospital. Pregnant patients are eligible to join the study through 28 weeks of gestation. Because the baseline risk of PTB is higher in multiple gestation pregnancies [37], only singleton pregnancies are eligible for CLEAR. However, not all pregnant patients know they are expecting multiples at the time of study recruitment; participants who consent to the study but are later determined to have multiple gestations are removed from the study.

Potentially eligible pregnant patients are identified through the HFH electronic medical record (EMR) system, Epic (Epic Systems Corp). Eligible patients are entered into a study-specific web-based database. This database is an internally developed HFH platform used for recruitment tracking, study communications, appointment scheduling, sample tracking, and incentive management.

CLEAR recruitment uses multiple approaches in an effort to be broadly appealing and accessible to all eligible patients. The primary points of introduction to the study are the following: (1) in-person recruitment at a prenatal care appointment, (2) a recruitment message via the HFH

online patient portal (MyChart; Epic Systems Corp), and (3) a mailed postal letter. Eligible patients may be contacted up to 8 additional times using a combination of texting, telephone calls, and email over a 6-week period or until they exit the eligibility window. Patients who decline participation or cannot be reached are assigned a final disposition code (eg, declined to participate, with a brief reason if available, or no contact, respectively) in the CLEAR database, after which no further contact is made.

Patients who express interest and are deemed eligible are sent a consent and Health Insurance Portability and Accountability Act (HIPAA) authorization form by email through the CLEAR Research Electronic Data Capture (REDCap) database. Study data are collected and managed using REDCap electronic data capture tools hosted at HFH [38,39]. REDCap is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources. Each participant is assigned the same study ID in both the CLEAR REDCap and the CLEAR web-based databases. Upon electronic signature of the CLEAR consent and HIPAA form, participants are automatically emailed a copy of the consent and HIPAA authorization form. CLEAR staff is simultaneously notified to review the form for completeness and update the participant’s status in the CLEAR web database to indicate enrollment.

Ethical Considerations

This study was approved by the HFH Institutional Review Board (13198 and 13306). Eligible pregnant participants

provide written informed consent and HIPAA authorization for both their and their baby’s participation. The consent is completed electronically via REDCap, which has previously been shown to be feasible and compliant in clinical research [40]. Participants are provided incentives (in the form of a gift card) for the study activities that they complete. Appropriate data use agreements and material transfer agreements have been obtained to share data and biospecimens across CLEAR projects and cores.

Data and Sample Collection

Surveys

After consenting, study questionnaires are sent via email or text, based on the participant’s GA, through REDCap (Figure 2). Surveys are administered up to 3 times during pregnancy: immediately after consent (baseline), in midpregnancy, and in late pregnancy (Figure 2). Participants entering the study “late” (after 22 weeks of gestation) complete the baseline survey and a combined mid- and late-pregnancy survey. Participants are allotted 2 weeks to complete each survey; deadlines have been shown to increase response rates in other settings [41,42]. If a participant delivers before completing the late pregnancy survey, an “early delivery” survey is sent that includes questions from the late pregnancy survey that are unlikely to be influenced by the early delivery (eg, family history of disease; Table 2). The early delivery survey is manually triggered by study staff and sent to the participant based on whether the participant and newborn have been discharged from the hospital or neonatal intensive care unit to avoid overburdening the family during a potentially stressful time.

Figure 2. Overview of the study activities based on the gestational age at which a participant provides informed consent. Study activities for participants who provided informed consent at <16 weeks are shown in pink, those who provided informed consent at 16 to 22 weeks are shown in yellow, and those who provided informed consent at >22 to 28 weeks are shown in purple.

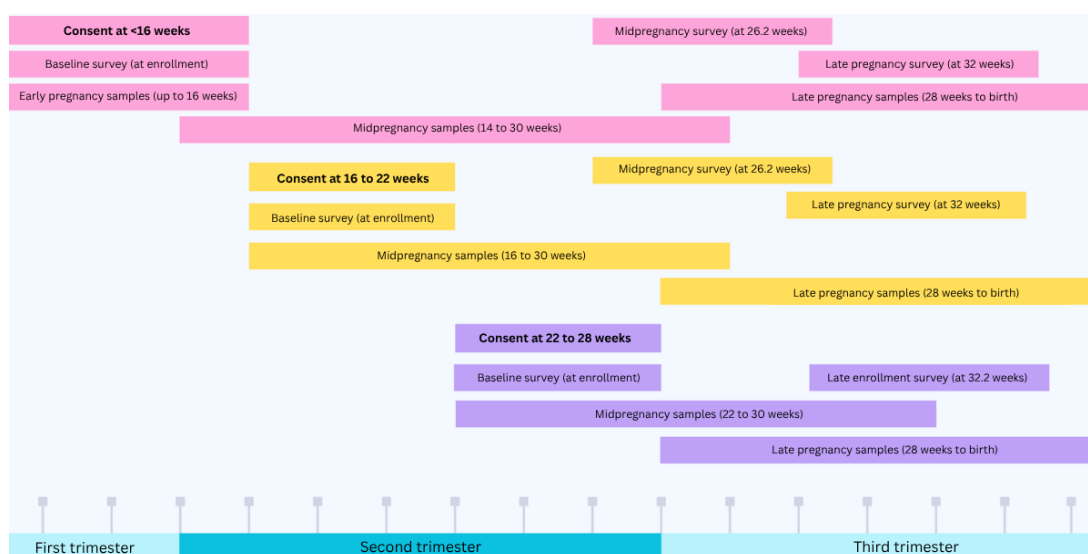


Table 2. Survey and electronic medical record data planned to be collected as part of the Center for Leadership in Environmental Awareness and Research study.

Data elements	Surveys					Electronic medical record
	Baseline	Midpregnancy	Late pregnancy	Late enrollment ^a	Early delivery ^b	
Maternal demographics						
Date of birth						✓
Race and ethnicity	✓					✓
Education	✓					
Employment status	✓					
Occupation	✓					
Marital status	✓					
Place of birth	✓					
Health insurance	✓					✓
Financial resources	✓					
Food insecurity	✓					
Address information (frequently visited locations)			✓	✓		✓
Prenatal health						
Nausea and vomiting	✓		✓			
Blood pressure						✓
Prepregnancy weight	✓					✓
Weight						✓
Height	✓					✓
Medications and supplements used			✓	✓		✓
Dental health			✓	✓		
Infections and illnesses			✓	✓	✓	✓
Maternal medical history						
Chronic conditions						✓
Weight at birth	✓					
Previous pregnancies						
Parity						✓
Gravidity						✓
Previous preterm birth						✓
History of pregnancy complications						✓
Maternal lifestyle						
Physical activity [43]	✓	✓	✓	✓		
Smoking or tobacco use			✓	✓	✓	✓
Alcohol use			✓	✓	✓	✓
Sleep		✓		✓		
Personal care products			✓	✓	✓	
Dietary screener [44,45]			✓	✓		
Pet ownership	✓					
Maternal psychosocial health						
Edinburgh Postnatal Depression Scale [46]	✓ ^c	✓	✓	✓		
Perceived Stress Scale [47]	✓	✓	✓	✓		
Confusion, Hubbub, and Order Scale [48]	✓					
PROMIS ^d Global Physical and Mental Health [49]	✓		✓	✓		
PROMIS Depression Short Form [50]	✓		✓	✓		
Brief Resilience Scale [51]	✓					

Data elements	Surveys					Electronic medical record
	Baseline	Midpregnancy	Late pregnancy	Late enrollment ^a	Early delivery ^b	
Connor-Davidson Resilience Scale [52]	✓					
Broad Autism Phenotype [53]	✓					
PROMIS Informational Support		✓		✓		
PROMIS Emotional Support [54]		✓		✓		
PROMIS Instrumental Support		✓		✓		
Household factors						
Type of housing	✓					
Basement	✓					
Number of inhabitants	✓					
Cleaning habits	✓					
Mold and mildew	✓					
Heating and cooling	✓					
Fresh air	✓					
Air purifiers or filters	✓					
Paternal demographics						
Date of birth	✓					
Race and ethnicity	✓					
Education	✓					
Location of birth	✓					
Paternal lifestyle and health history						
Height	✓					
Weight	✓					
Smoking	✓					
Baby's family health history			✓	✓	✓	
Delivery information						
Date of delivery						✓
Gestational age at delivery						✓
Labor induction						✓
Labor augmentation						✓
Mode of delivery						✓
Complications						✓
Medications used						✓

^aReplaces mid and late pregnancy surveys if participant enrolls after 32.2 weeks' gestation.

^bReplaces the late pregnancy survey if participant delivers before completing it.

^cAnxiety subscale only [55].

^dPROMIS: Patient-Reported Outcomes Measurement Information System.

EMR Data

The EMR is used to obtain maternal medical history, including vitals (blood pressure and weight) over pregnancy, medication use over pregnancy, PTB interventions (eg, cerclage), and delivery information including date of delivery and GA at delivery (Table 2).

Biological Specimens

Overview

Biological samples are collected prenatally (up to 3 times during pregnancy; Figure 2 and Table 3), and after delivery. Table 3 lists the biological samples collected, their frequency,

and a brief reason for each specimen type. Participants can refuse collection of any biospecimen at any time point and remain in the study.

Each biological sample type is assigned a unique ID number that indicates study ID information as well as visit number, sample type, and aliquot number. No personal identifying information is included in these IDs. Each sample ID is printed as a barcode label that is attached to all specimen collection containers and aliquot containers. All samples and aliquots are scanned into the CLEAR biospecimen database along with their storage location. This database is also used for tracking shipment of specimens for analyses.

Table 3. Biological samples collected from pregnant participants in the Center for Leadership in Environmental Awareness and Research study.

Sample type	Time points	Primary purpose
Urine	Up to 3 times during pregnancy (early, mid, and late pregnancy)	Measurement of volatile organic compound metabolites
Blood (serum and plasma)	Up to 3 times during pregnancy (early, mid, and late pregnancy)	Measurement of inflammatory biomarkers
Placenta	After delivery	Epigenetic and transcriptomic analyses

Urine

Participants are asked to provide a spot urine sample during a research clinic visit, up to 3 times during pregnancy (early, mid, and late pregnancy visits; Table 3). Urine samples are immediately refrigerated and brought to the HFH Molecular Epidemiology Research Laboratory (MERL). The urine specimen is gently mixed, and urine specific gravity measured with a refractometer before 5 aliquots, each with 1800 μL of urine, are created and stored at -80°C . VOC metabolite measurement has successfully been completed in urine samples in previous studies during pregnancy [56, 57], including in our own preliminary work [23]. The VOC metabolites planned to be measured in CLEAR are listed in Table 1.

Blood

At research clinic visits, pregnant participants are asked to provide a blood sample, up to 3 times during pregnancy (early, mid, and late pregnancy visits; Table 3). Trained phlebotomists obtain the blood samples (1 ethylene diamine tetraacetic acid and 1 sodium heparin tube). Blood samples are brought back to the HFH MERL for processing. The sodium heparin tube is gently inverted several times and centrifuged at $1300\times g$ at room temperature for 10 minutes. In total, 5 aliquots are created: the first containing 1000 μL of sample and the remaining 4 aliquots each containing 500 μL of sample. Aliquots are stored at -80°C . The ethylene diamine tetraacetic acid tube is processed similarly, except that 1.5 mL of whole blood is aliquoted before centrifuging, and the buffy coat is retained after removal of plasma. Plasma will be used to measure inflammatory biomarkers during pregnancy.

Placenta

Placentas are collected after delivery in partnership with the HFH Departments of Pathology and WHS. After delivery of the placenta, WHS staff place the placenta in a placenta collection bucket labeled with the patient's medical record number, and the bucket is placed in a refrigerator on the labor and delivery floor. The refrigerator has shelves specifically designated for research and pathology. Placentas that require pathological evaluation are retrieved from the HFH Department of Pathology after clinical evaluation. As such, placentas retrieved from pathology are weighed, formalin fixed, and sectioned as part of the clinical evaluation. Placentas not intended for clinical evaluation by the Department of Pathology are picked up daily by the study team and transferred on ice to the MERL for fixation and processing. To the extent possible, processing of placentas that are not sent for clinical pathologic evaluation mirrors that of those sent for pathologic evaluation. This includes weight,

length, and thickness measurement, fixation in 10% buffered formalin, and gross evaluation by a histotechnologist. From all fixed placentas, 4 full-thickness sections (1 from each quadrant), 1 section of umbilical cord, and 1 membrane roll are paraffin embedded and stored in cassettes for downstream analyses (eg, DNA and RNA extraction).

Optional CLEAR Study Activities

As part of the consenting process, participants are asked to opt in or opt out of several optional study activities. Although the participant does have to make a choice to opt in or opt out for each optional study activity, their specific choice does not impact their overall participation in CLEAR. The optional study activities include the following:

- Receipt of text messages from the Twilio system (Twilio Inc), a secure and HIPAA-compliant messaging system that integrates within REDCap
- Collection and storage of leftover clinical samples (eg, urine and blood) from the (1) pregnant participant and (2) child participant
- Collection of newborn screening blood spots (Michigan BioTrust for Health [58])
- Foliage sampling (collaborative work with CLEAR engineering project 1)
- Consent to contact by the CLEAR Community Engagement Core for Healthy Homes (Indoor Home Safety Assessment [59])
- Long-term storage of data and samples for future research

Primary Health Outcome: PTB

EMR data will be used to determine GA at delivery and define PTB. The "minimal" criteria described by Pennell et al [60] for determining PTB status will be used. GA at delivery will be assigned in accordance with the American College of Obstetricians and Gynecologists guidelines [61]. Our primary analysis will focus on cases defined as PTB if GA at delivery <37 weeks. Secondary analyses will consider the following: (1) continuous GA at delivery, (2) PTB subtypes (eg, extremely preterm if <28 weeks, very preterm if 28 to 32 weeks, and moderate-to-late preterm if 32 to 37 weeks), and (3) whether the PTB was spontaneous (preterm premature rupture of membranes or preterm labor) or indicated for medical reasons.

Sample Size and Power

The CLEAR birth cohort was designed to examine the association of VOCs with PTB, to examine if VOCs impact risk of PTB via maternal inflammation and placental alterations, and to explore potential sources of VOC exposures. A nested case-control approach was chosen to

weigh the high cost of laboratory analyses against the need to recruit patients early enough in pregnancy to accrue repeated and multiple biospecimens; nested case-control designs reduce some costs while having minimal impact on statistical efficiency [62]. Given birth rates at HFH and past birth cohort recruitment [63], we anticipate recruiting 1075 pregnant participants into CLEAR. The PTB rate in Detroit in 2018 was 15.2% [64]. We assume that approximately 15% of the CLEAR cohort will experience PTB; thus, 164 PTB cases are expected. Controls with GA at delivery ≥ 37 weeks at delivery will be 1:1 frequency matched [65] to cases based on infant sex and maternal race. Frequency matching maintains power [66] yet allows us to still adjust for these variables in the analysis. This sample provides sufficient power for the proposed analyses. For example, based on data from pregnant women in the National Children's Study [56], SD of VOC metabolites on the \log_{10} scale is expected to range from 0.1 to 0.6. With 164 per group for the nested case-control, $\alpha=.05$, 2-tailed tests, and 80% power, an odds ratio of 1.3, 1.6, or 3.5 for a VOC metabolite with an SD of 0.6, 0.3, or 0.1, respectively, could be detected for associations of a VOC metabolite with PTB. For examining the association of VOCs with placental function, with a false discovery rate of 0.05, there is sufficient power ($>80\%$) to detect correlations of ≥ 0.3 and fold changes of 2.5.

General Statistical Considerations

A combined directed acyclic graph and change-in-estimate approach will guide selection of confounders and effect modifiers [67-70]. Although a large number of potential confounders will be measured, observational studies are subject to unmeasured confounding: e-values will be used to estimate the effect of unmeasured confounding [71,72]. We will consider methods to account for potential bias or misclassification of PTB case or control status by including the method of assigning GA at delivery (first-trimester ultrasound, second-trimester ultrasound, or date of last menstrual period) in models.

VOC results below the limit of quantification will be imputed as in Wei et al [73]. Multiple testing adjustment for high-dimensional data, such as false discovery rate, will be used [74]. With up to 6 VOCs measured, to reduce the multiple-testing burden, chemical mixture modeling will be used. Potential chemical mixture approaches include Bayesian kernel machine regression, weighted quantile sum regression, and principal components analysis [75]; each has distinct strengths and weaknesses. If more than one method is potentially appropriate, robustness of findings will be examined across different methods and all findings will be reported as in Vuong et al [76].

To examine associations between VOC metabolites and PTB, generalized linear mixed regression models will be used, where the repeated VOC metabolites are the covariates and PTB status (yes or no) is the outcome of interest. Generalized linear mixed model elastic net (glmmlasso), which accounts for mixtures and correlation between repeated measures, will be used to examine associations between VOC metabolites and maternal inflammation [77]. The association

of VOC mixtures with DNA methylation will be tested using generalized network-based elastic-net linear mixed model (GELMMnet) with a logit link for β values (equivalent to M values), and associations with gene expression (RNAseq) will be tested using GELMMnet with a Poisson distribution [78]. If significant associations are identified above, we will then test if inflammation and/or placental function mediates associations between VOC metabolites and PTB using a mediation approach for dichotomous outcomes that estimates the natural direct and indirect effects for each individual marker [79], which has been extended to account for repeated measures and high-dimensional mediation effects [80]. To explore neighborhood-level or related factors, clustering or hot spot analysis and geographically weighted regression, a spatial analysis technique that uses nonstationary data (eg, demographic factors and environmental characteristics) to model the local relationship between predictors and an outcome (ie, VOC level), will be used [81,82].

Engagement and Retention

Multiple strategies have been implemented to promote participant engagement and retention throughout the study. Participants receive personalized communications, including birthday and welcome baby cards, to acknowledge important milestones and maintain an ongoing connection with the study. CLEAR's partnerships with clinical providers support trust-building and continuity, while study staff are trained to create positive, supportive encounters that help participants feel valued and respected. To help reduce barriers to continued participation, logistical support is offered, including Lyft (Lyft Inc) rides to and from study appointments, parking reimbursement, and a resource flyer tailored to participant needs (eg, mental health and postpartum assistance). Participants also receive prenatal gifts (eg, diaper bag and vanity kit) and age-appropriate gifts for their baby as tokens of appreciation. Collectively, these strategies are designed to foster a supportive study environment, reduce participant burden, and enhance long-term retention.

The consent process incorporates language specific to follow-up, including permission to use publicly available directories, databases (eg, WhitePages), or social media (eg, Facebook; Meta Platforms Inc) to validate contact information if participants are difficult to reach. This ensures that participants are informed of, and comfortable with, strategies to support continued engagement over time. In addition, participants are asked to provide the name and contact information of a family member or close friend that the CLEAR study can reach out to if the participant is lost to follow-up.

Results

Funding for the study began September 2022. Study recruitment began in November 2023. As of April 22, 2026, a total of 468 pregnant patients have consented to participate in the CLEAR birth cohort. The mean age at consent was 29.9 (SD 6.1) years. On the basis of race information captured in the EMR, most consented participants are Black (346/468,

73.9%), and 16% (75/468) are White, with the remainder reporting either another race, declining to provide race information, or having unknown race. Delivery information is available for 307 (65.6%) participants; of these, 13.7% (42/307) were PTB.

Recruitment is ongoing, with 43.5% (468/1075) of the recruitment goal achieved. Continued efforts to increase recruitment and retention rates will be used, as well as potential expansion of recruitment at other Detroit-based WHS clinics.

Discussion

Anticipated Findings

Establishment and study of the CLEAR birth cohort will provide needed data on the role of VOCs in PTB, particularly in older, urban areas such as Detroit, Michigan. Although previous studies have suggested potential associations between VOCs and PTB or other adverse birth outcomes, nearly all have relied on VOC exposures estimated based on residential address, with a single study examining personal exposure to only a single VOC (ie, benzene) [5]. CLEAR will address this limitation via measurement of VOC metabolites in maternal urine across pregnancy. Additionally, better understanding of potential mechanistic pathways, such as impacts on maternal inflammation and on placental function, may reveal areas for interventions to prevent PTB. To date, limited studies have examined these relationships with direct measurements of VOCs and/or in placental tissue; thus, CLEAR will provide unique data on these potential associations. With detailed address information, the CLEAR birth cohort will also allow for the study of potential VOC exposure sources, which are important to identify ways to either decrease exposure or to remediate contaminated areas. Previous studies have identified sources of indoor VOC levels in the homes of children with asthma in Detroit [83]; CLEAR will expand on these previous findings by better understanding sources of direct VOC measurements in pregnant women.

The CLEAR birth cohort addresses limitations of prior studies, importantly by directly measuring VOC exposure via urinary metabolites across pregnancy. Potential limitations, inherent to observational studies, remain. Although CLEAR captures numerous measures of potential confounders, residual confounding remains a potential limitation. This is somewhat mitigated by the proposed use of e-values. Similarly, recruitment bias is a potential issue impacting clinical research studies; because we are recruiting through a health system, we can describe those who do and do not agree to participate in the CLEAR birth cohort (in terms of basic demographic characteristics) which will enable us to address potential bias. Although we propose to measure urinary metabolites and circulating inflammatory biomarkers across pregnancy, not all women enter prenatal care in the first trimester, nor do they complete all study visits, thus missing data are anticipated. Similarly, women who do not deliver at a study hospital or who wish to retain their placenta for personal reasons may not have a placenta for functional measurements. Imputation and/or methods that account for loss-to-follow-up such as inverse probability weighting will be used when feasible. As a postindustrial rust belt city, results will be most generalizable to other similar urban cities, most of which have similarly high PTB rates as Detroit [1,2].

Future Directions

Establishment of the CLEAR birth cohort, currently focused on PTB, provides opportunities for additional future maternal and child research studies. Although outside the scope of the primary aims of this study, additional maternal and infant data collections are planned, including collection of maternal hair, toenails, and stool for assessment of stress biomarkers, additional environmental exposures, and microbiota, respectively, and neonatal blood spots, urine, and stool for measures of inflammation, metabolomics, and microbiota, respectively. Additional surveys over infancy are also planned to measure other exposures, child health and neurodevelopment, and other lifestyle factors.

Acknowledgments

The authors convey special thanks to Dustyn Levenson for designing Figure 1. The authors also acknowledge the important contributions of the pilot participants and the participants of the Center for Leadership in Environmental Awareness and Research birth cohort; without them, none of this work would be possible.

Funding

This study was supported by the National Institutes of Health (grants P42 ES030991 and P01 AI089473).

Data Availability

Given the sensitive nature of the Center for Leadership in Environmental Awareness and Research birth cohort data, most data will be available upon request and with appropriate institutional review board and data use agreements. Limited datasets will be deposited into relevant repositories following submission and publication of the main manuscripts and in accordance with National Institutes of Health guidelines. Data collection is ongoing, and there are currently no available data to share.

Authors' Contributions

Conceptualization: AEC-B, JKS

Funding acquisition: AEC-B, JKS

Investigation: AEC-B, EL, AML, ML, JKS

Methodology: AEC-B, EL, AML, ML, JKS

Project administration: AEC-B, EL, JKS
Resources: AEC-B, JKS
Supervision: AEC-B, JKS
Visualization: JKS, EL (lead), DL (supporting)
Writing—original draft: AEC-B, EL, JKS
Writing—review and editing: AEC-B, EL, AML, ML, JKS

Conflicts of Interest

None declared.

Checklist 1

STROBE checklist.

[\[DOCX File \(Microsoft Word File\), 35 KB-Checklist 1\]](#)

References

- 2024 March of Dimes report card: the state of maternal and infant health for American families. March of Dimes. 2024. URL: https://www.marchofdimes.org/sites/default/files/2024-11/2024_MOD_Report_Card_and_Policy_Actions_Booklet_V1.pdf [Accessed 2025-10-24]
- State summary for Michigan. March of Dimes Peristats. URL: <https://www.marchofdimes.org/peristats/state-summaries/michigan?top=3&lev=1&stop=55@=99&sreg=26&creg=2622000&obj=8&slev=5> [Accessed 2025-10-24]
- Lougheed T. Arising from the ashes? Environmental health in Detroit. *Environ Health Perspect*. Dec 2014;122(12):A324-31. [doi: [10.1289/ehp.122-A324](https://doi.org/10.1289/ehp.122-A324)] [Medline: [25438340](https://pubmed.ncbi.nlm.nih.gov/25438340/)]
- Miller CJ, Runge-Morris M, Cassidy-Bushrow AE, et al. A review of volatile organic compound contamination in post-industrial urban centers: reproductive health implications using a Detroit lens. *Int J Environ Res Public Health*. Nov 25, 2020;17(23):8755. [doi: [10.3390/ijerph17238755](https://doi.org/10.3390/ijerph17238755)] [Medline: [33255777](https://pubmed.ncbi.nlm.nih.gov/33255777/)]
- Slama R, Thiebaugeorges O, Goua V, et al. Maternal personal exposure to airborne benzene and intrauterine growth. *Environ Health Perspect*. Aug 2009;117(8):1313-1321. [doi: [10.1289/ehp.0800465](https://doi.org/10.1289/ehp.0800465)] [Medline: [19672414](https://pubmed.ncbi.nlm.nih.gov/19672414/)]
- Serrano-Lomelin J, Nielsen CC, Jabbar MS, et al. Interdisciplinary-driven hypotheses on spatial associations of mixtures of industrial air pollutants with adverse birth outcomes. *Environ Int*. Oct 2019;131:104972. [doi: [10.1016/j.envint.2019.104972](https://doi.org/10.1016/j.envint.2019.104972)] [Medline: [31299602](https://pubmed.ncbi.nlm.nih.gov/31299602/)]
- Santos DA, Nascimento LF. Maternal exposure to benzene and toluene and preterm birth. A longitudinal study. *Sao Paulo Med J*. 2019;137(6):486-490. [doi: [10.1590/1516-3180.2019.0224170919](https://doi.org/10.1590/1516-3180.2019.0224170919)] [Medline: [32159633](https://pubmed.ncbi.nlm.nih.gov/32159633/)]
- Llop S, Ballester F, Estarlich M, Esplugues A, Rebagliato M, Iñiguez C. Preterm birth and exposure to air pollutants during pregnancy. *Environ Res*. Nov 2010;110(8):778-785. [doi: [10.1016/j.envres.2010.09.009](https://doi.org/10.1016/j.envres.2010.09.009)] [Medline: [20932516](https://pubmed.ncbi.nlm.nih.gov/20932516/)]
- Estarlich M, Ballester F, Davdand P, et al. Exposure to ambient air pollution during pregnancy and preterm birth: a Spanish multicenter birth cohort study. *Environ Res*. May 2016;147:50-58. [doi: [10.1016/j.envres.2016.01.037](https://doi.org/10.1016/j.envres.2016.01.037)] [Medline: [26851724](https://pubmed.ncbi.nlm.nih.gov/26851724/)]
- Wilhelm M, Ghosh JK, Su J, Cockburn M, Jerrett M, Ritz B. Traffic-related air toxics and preterm birth: a population-based case-control study in Los Angeles County, California. *Environ Health*. Oct 7, 2011;10(1):89. [doi: [10.1186/1476-069X-10-89](https://doi.org/10.1186/1476-069X-10-89)] [Medline: [21981989](https://pubmed.ncbi.nlm.nih.gov/21981989/)]
- Ruckart PZ, Bove FJ, Maslia M. Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: a cross-sectional study. *Environ Health*. Nov 20, 2014;13(1):99. [doi: [10.1186/1476-069X-13-99](https://doi.org/10.1186/1476-069X-13-99)] [Medline: [25413571](https://pubmed.ncbi.nlm.nih.gov/25413571/)]
- Poirier A, Dodds L, Dummer T, Rainham D, Maguire B, Johnson M. Maternal exposure to air pollution and adverse birth outcomes in Halifax, Nova Scotia. *J Occup Environ Med*. Dec 2015;57(12):1291-1298. [doi: [10.1097/JOM.0000000000000604](https://doi.org/10.1097/JOM.0000000000000604)] [Medline: [26641824](https://pubmed.ncbi.nlm.nih.gov/26641824/)]
- Keeler C, Luben TJ, Forestieri N, Olshan AF, Desrosiers TA. Is residential proximity to polluted sites during pregnancy associated with preterm birth or low birth weight? Results from an integrated exposure database in North Carolina (2003-2015). *J Expo Sci Environ Epidemiol*. Mar 2023;33(2):229-236. [doi: [10.1038/s41370-022-00475-0](https://doi.org/10.1038/s41370-022-00475-0)] [Medline: [36100666](https://pubmed.ncbi.nlm.nih.gov/36100666/)]
- Cassidy-Bushrow AE, Burmeister C, Lamerato L, et al. Prenatal airshed pollutants and preterm birth in an observational birth cohort study in Detroit, Michigan, USA. *Environ Res*. Oct 2020;189:109845. [doi: [10.1016/j.envres.2020.109845](https://doi.org/10.1016/j.envres.2020.109845)] [Medline: [32678729](https://pubmed.ncbi.nlm.nih.gov/32678729/)]
- Minciullo PL, Navarra M, Calapai G, Gangemi S. Cytokine network involvement in subjects exposed to benzene. *J Immunol Res*. 2014;2014:937987. [doi: [10.1155/2014/937987](https://doi.org/10.1155/2014/937987)] [Medline: [25202711](https://pubmed.ncbi.nlm.nih.gov/25202711/)]

16. Kido T, Sugaya C, Ikeuchi R, Kudo Y, Tsunoda M, Aizawa Y. The increases in mRNA expressions of inflammatory cytokines by adding cleaning solvent or tetrachloroethylene in the Murine macrophage cell line J774.1 evaluated by real-time PCR. *Ind Health*. 2013;51(3):319-325. [doi: [10.2486/indhealth.ms1364](https://doi.org/10.2486/indhealth.ms1364)] [Medline: [23538726](https://pubmed.ncbi.nlm.nih.gov/23538726/)]
17. Fenga C, Gangemi S, Costa C. Benzene exposure is associated with epigenetic changes (review). *Mol Med Rep*. Apr 2016;13(4):3401-3405. [doi: [10.3892/mmr.2016.4955](https://doi.org/10.3892/mmr.2016.4955)] [Medline: [26936331](https://pubmed.ncbi.nlm.nih.gov/26936331/)]
18. Blossom SJ, Gilbert KM. Epigenetic underpinnings of developmental immunotoxicity and autoimmune disease. *Curr Opin Toxicol*. Aug 2018;10:23-30. [doi: [10.1016/j.cotox.2017.11.013](https://doi.org/10.1016/j.cotox.2017.11.013)] [Medline: [30613805](https://pubmed.ncbi.nlm.nih.gov/30613805/)]
19. Thomas R, McHale CM, Lan Q, et al. Global gene expression response of a population exposed to benzene: a pilot study exploring the use of RNA-sequencing technology. *Environ Mol Mutagen*. Aug 2013;54(7):566-573. [doi: [10.1002/em.21801](https://doi.org/10.1002/em.21801)] [Medline: [23907980](https://pubmed.ncbi.nlm.nih.gov/23907980/)]
20. Hong JY, Yu SY, Kim GW, et al. Identification of time-dependent biomarkers and effects of exposure to volatile organic compounds using high-throughput analysis. *Environ Toxicol*. Nov 2016;31(11):1563-1570. [doi: [10.1002/tox.22160](https://doi.org/10.1002/tox.22160)] [Medline: [26018793](https://pubmed.ncbi.nlm.nih.gov/26018793/)]
21. Schiffman C, McHale CM, Hubbard AE, et al. Identification of gene expression predictors of occupational benzene exposure. *PLoS ONE*. 2018;13(10):e0205427. [doi: [10.1371/journal.pone.0205427](https://doi.org/10.1371/journal.pone.0205427)] [Medline: [30300410](https://pubmed.ncbi.nlm.nih.gov/30300410/)]
22. Phillips RV, Rieswijk L, Hubbard AE, et al. Human exposure to trichloroethylene is associated with increased variability of blood DNA methylation that is enriched in genes and pathways related to autoimmune disease and cancer. *Epigenetics*. Nov 2019;14(11):1112-1124. [doi: [10.1080/15592294.2019.1633866](https://doi.org/10.1080/15592294.2019.1633866)] [Medline: [31241004](https://pubmed.ncbi.nlm.nih.gov/31241004/)]
23. Cassidy-Bushrow AE, Burmeister C, Birbeck J, et al. Ambient BTEX exposure and mid-pregnancy inflammatory biomarkers in pregnant African American women. *J Reprod Immunol*. Jun 2021;145:103305. [doi: [10.1016/j.jri.2021.103305](https://doi.org/10.1016/j.jri.2021.103305)] [Medline: [33725526](https://pubmed.ncbi.nlm.nih.gov/33725526/)]
24. Straughen JK, Loveless I, Chen Y, et al. The impact of environmental benzene, toluene, ethylbenzene, and xylene exposure on blood-based DNA methylation profiles in pregnant African American women from Detroit. *Int J Environ Res Public Health*. Feb 23, 2024;21(3):256. [doi: [10.3390/ijerph21030256](https://doi.org/10.3390/ijerph21030256)] [Medline: [38541258](https://pubmed.ncbi.nlm.nih.gov/38541258/)]
25. Lehmann I, Thoeleke A, Rehwagen M, et al. The influence of maternal exposure to volatile organic compounds on the cytokine secretion profile of neonatal T cells. *Environ Toxicol*. 2002;17(3):203-210. [doi: [10.1002/tox.10055](https://doi.org/10.1002/tox.10055)] [Medline: [12112628](https://pubmed.ncbi.nlm.nih.gov/12112628/)]
26. Baiz N, Slama R, Béné MC, et al. Maternal exposure to air pollution before and during pregnancy related to changes in newborn's cord blood lymphocyte subpopulations. The EDEN study cohort. *BMC Pregnancy Childbirth*. Nov 2, 2011;11(87):87. [doi: [10.1186/1471-2393-11-87](https://doi.org/10.1186/1471-2393-11-87)] [Medline: [22047167](https://pubmed.ncbi.nlm.nih.gov/22047167/)]
27. Toxicological profile for trichloroethylene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services. 2025. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp19.pdf> [Accessed 2025-05-25]
28. Toxicological profile for tetrachloroethylene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services. 2025. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp18.pdf> [Accessed 2025-05-25]
29. Toxicological profile for benzene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services. 2024. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp3.pdf> [Accessed 2025-05-25]
30. Toxicological profile for toluene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services, Public Health Service. 2017. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp56.pdf> [Accessed 2025-05-25]
31. Toxicological profile for ethylbenzene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services, Public Health Service. 2010. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp110.pdf> [Accessed 2025-05-25]
32. Ethylbenzene (100-41-4). Environmental Protection Agency. 1992. URL: <https://www.epa.gov/sites/default/files/2016-09/documents/ethylbenzene.pdf> [Accessed 2025-05-25]
33. Toxicological profile for xylene. Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services, Public Health Service. 2007. URL: <https://www.atsdr.cdc.gov/toxprofiles/tp71.pdf> [Accessed 2025-05-25]
34. Lyons NA, Redding A, Susick LL, Leydet EM, Tyra MA, Santarossa S. Using new age recruitment methods. *Int J Soc Media Online Communities*. Jan 2023;15(1):1-15. [doi: [10.4018/IJSMOC.324057](https://doi.org/10.4018/IJSMOC.324057)]
35. Santarossa S, Redding A, Connell M, Kao K, Susick L, Kerver JM. Exploring preliminary dietary intake results using a novel dietary assessment tool with pregnant participants enrolled in a birth cohort. *BMC Res Notes*. Feb 1, 2024;17(1):42. [doi: [10.1186/s13104-024-06697-9](https://doi.org/10.1186/s13104-024-06697-9)] [Medline: [38303032](https://pubmed.ncbi.nlm.nih.gov/38303032/)]
36. Redding A, Coates D, Cassidy-Bushrow AE, Straughen JK, Santarossa S. Activity tracking devices in pregnancy: understanding the participant experience in a longitudinal birth cohort. *Womens Health (Lond)*. 2025;21:17455057251344388. [doi: [10.1177/17455057251344388](https://doi.org/10.1177/17455057251344388)] [Medline: [40616530](https://pubmed.ncbi.nlm.nih.gov/40616530/)]

37. Fuchs F, Senat MV. Multiple gestations and preterm birth. *Semin Fetal Neonatal Med.* Apr 2016;21(2):113-120. [doi: [10.1016/j.siny.2015.12.010](https://doi.org/10.1016/j.siny.2015.12.010)] [Medline: [26795885](https://pubmed.ncbi.nlm.nih.gov/26795885/)]
38. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform.* Jul 2019;95:103208. [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* Apr 2009;42(2):377-381. [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
40. Frelich MJ, Bosler ME, Gould JC. Research Electronic Data Capture (REDCap) electronic Informed Consent Form (eICF) is compliant and feasible in a clinical research setting. *Int J Clin Trials.* 2015;2(3):51. [doi: [10.18203/2349-3259.ijct20150591](https://doi.org/10.18203/2349-3259.ijct20150591)]
41. Décieux JP, Heinz A. Does a short-term deadline extension affect participation rates of an online survey? Experimental evidence from an online panel. *Int J Soc Res Methodol.* May 3, 2024;27(3):369-373. [doi: [10.1080/13645579.2022.2153475](https://doi.org/10.1080/13645579.2022.2153475)]
42. Martin E. Can a deadline and compressed mailing schedule improve mail response in the decennial census? *Public Opin Q.* 2009;73(2):361-367. [doi: [10.1093/poq/nfp018](https://doi.org/10.1093/poq/nfp018)]
43. Chasan-Taber L, Schmidt MD, Roberts DE, Hosmer D, Markenson G, Freedson PS. Development and validation of a pregnancy physical activity questionnaire. *Med Sci Sports Exerc.* Oct 2004;36(10):1750-1760. [doi: [10.1249/01.mss.0000142303.49306.0d](https://doi.org/10.1249/01.mss.0000142303.49306.0d)] [Medline: [15595297](https://pubmed.ncbi.nlm.nih.gov/15595297/)]
44. Dietary Screener Questionnaire (DSQ) in the NHANES 2009-10: dietary factors, food items asked, and testing status for DSQ. Division of Cancer Control and Population Sciences, National Cancer Institute, National Institute of Health. 2025. URL: <https://epi.grants.cancer.gov/nhanes/dietscreen/evaluation.html> [Accessed 2025-05-25]
45. Thompson FE, Midthune D, Kahle L, Dodd KW. Development and evaluation of the National Cancer Institute's dietary screener questionnaire scoring algorithms. *J Nutr.* Jun 2017;147(6):1226-1233. [doi: [10.3945/jn.116.246058](https://doi.org/10.3945/jn.116.246058)] [Medline: [28490673](https://pubmed.ncbi.nlm.nih.gov/28490673/)]
46. Cox JL, Chapman G, Murray D, Jones P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *J Affect Disord.* Jul 29, 1996;39(3):185-189. [doi: [10.1016/0165-0327\(96\)00008-0](https://doi.org/10.1016/0165-0327(96)00008-0)] [Medline: [8856422](https://pubmed.ncbi.nlm.nih.gov/8856422/)]
47. Cohen S, Williamson G. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. *The Social Psychology of Health: Claremont Symposium on Applied Social Psychology.* Sage; 1988. URL: <https://archive.org/details/socialpsychology0000clar/page/n7/mode/2up> [Accessed 2025-05-25] ISBN: 9780803931626
48. Matheny AP, Wachs TD, Ludwig JL, Phillips K. Bringing order out of chaos: psychometric characteristics of the confusion, hubbub, and order scale. *J Appl Dev Psychol.* Jul 1995;16(3):429-444. [doi: [10.1016/0193-3973\(95\)90028-4](https://doi.org/10.1016/0193-3973(95)90028-4)]
49. Hays RD, Schalet BD, Spritzer KL, Cella D. Two-item PROMIS® global physical and mental health scales. *J Patient Rep Outcomes.* 2017;1(1):2. [doi: [10.1186/s41687-017-0003-8](https://doi.org/10.1186/s41687-017-0003-8)] [Medline: [29757325](https://pubmed.ncbi.nlm.nih.gov/29757325/)]
50. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment.* Sep 2011;18(3):263-283. [doi: [10.1177/1073191111411667](https://doi.org/10.1177/1073191111411667)] [Medline: [21697139](https://pubmed.ncbi.nlm.nih.gov/21697139/)]
51. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. *Int J Behav Med.* 2008;15(3):194-200. [doi: [10.1080/10705500802222972](https://doi.org/10.1080/10705500802222972)] [Medline: [18696313](https://pubmed.ncbi.nlm.nih.gov/18696313/)]
52. Connor KM, Davidson JR. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depress Anxiety.* 2003;18(2):76-82. [doi: [10.1002/da.10113](https://doi.org/10.1002/da.10113)] [Medline: [12964174](https://pubmed.ncbi.nlm.nih.gov/12964174/)]
53. Hurley RS, Losh M, Parlier M, Reznick JS, Piven J. The broad autism phenotype questionnaire. *J Autism Dev Disord.* Oct 2007;37(9):1679-1690. [doi: [10.1007/s10803-006-0299-3](https://doi.org/10.1007/s10803-006-0299-3)] [Medline: [17146701](https://pubmed.ncbi.nlm.nih.gov/17146701/)]
54. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol.* Nov 2010;63(11):1179-1194. [doi: [10.1016/j.jclinepi.2010.04.011](https://doi.org/10.1016/j.jclinepi.2010.04.011)] [Medline: [20685078](https://pubmed.ncbi.nlm.nih.gov/20685078/)]
55. Phillips J, Charles M, Sharpe L, Matthey S. Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants. *J Affect Disord.* Nov 2009;118(1-3):101-112. [doi: [10.1016/j.jad.2009.02.004](https://doi.org/10.1016/j.jad.2009.02.004)] [Medline: [19275960](https://pubmed.ncbi.nlm.nih.gov/19275960/)]
56. Boyle EB, Viet SM, Wright DJ, et al. Assessment of exposure to VOCs among pregnant women in the National Children's Study. *Int J Environ Res Public Health.* Mar 29, 2016;13(4):376. [doi: [10.3390/ijerph13040376](https://doi.org/10.3390/ijerph13040376)] [Medline: [27043585](https://pubmed.ncbi.nlm.nih.gov/27043585/)]
57. Li M, Wan Y, Qian X, et al. Urinary metabolites of multiple volatile organic compounds among pregnant women across pregnancy: variability, exposure characteristics, and associations with selected oxidative stress biomarkers. *Environ Int.* Mar 2023;173:107816. [doi: [10.1016/j.envint.2023.107816](https://doi.org/10.1016/j.envint.2023.107816)] [Medline: [36805810](https://pubmed.ncbi.nlm.nih.gov/36805810/)]

58. Langbo C, Bach J, Kleyn M, Downes FP. From newborn screening to population health research: implementation of the Michigan BioTrust for health. *Public Health Rep.* 2013;128(5):377-384. [doi: [10.1177/003335491312800508](https://doi.org/10.1177/003335491312800508)] [Medline: [23997284](https://pubmed.ncbi.nlm.nih.gov/23997284/)]
59. Larson PS, Gronlund C, Thompson L, et al. Recurrent home flooding in Detroit, MI 2012-2020: results of a household survey. *Int J Environ Res Public Health.* Jul 19, 2021;18(14):7659. [doi: [10.3390/ijerph18147659](https://doi.org/10.3390/ijerph18147659)] [Medline: [34300113](https://pubmed.ncbi.nlm.nih.gov/34300113/)]
60. Pennell CE, Jacobsson B, Williams SM, et al. Genetic epidemiologic studies of preterm birth: guidelines for research. *Am J Obstet Gynecol.* Feb 2007;196(2):107-118. [doi: [10.1016/j.ajog.2006.03.109](https://doi.org/10.1016/j.ajog.2006.03.109)] [Medline: [17306646](https://pubmed.ncbi.nlm.nih.gov/17306646/)]
61. Committee opinion no 700: methods for estimating the due date. *Obstet Gynecol.* May 2017;129(5):e150-e154. [doi: [10.1097/AOG.0000000000002046](https://doi.org/10.1097/AOG.0000000000002046)] [Medline: [28426621](https://pubmed.ncbi.nlm.nih.gov/28426621/)]
62. Ernster VL. Nested case-control studies. *Prev Med.* Sep 1994;23(5):587-590. [doi: [10.1006/pmed.1994.1093](https://doi.org/10.1006/pmed.1994.1093)] [Medline: [7845919](https://pubmed.ncbi.nlm.nih.gov/7845919/)]
63. Cassidy-Bushrow AE, Peters RM, Johnson DA, Li J, Rao DS. Vitamin D nutritional status and antenatal depressive symptoms in African American women. *J Womens Health (Larchmt).* Nov 2012;21(11):1189-1195. [doi: [10.1089/jwh.2012.3528](https://doi.org/10.1089/jwh.2012.3528)] [Medline: [22823176](https://pubmed.ncbi.nlm.nih.gov/22823176/)]
64. Preterm birth rate: Detroit and Michigan, 2018-2024. *March of Dimes Peristats.* 2025. URL: <https://www.marchofdimes.org/peristats/data?reg=99&top=3&stop=60&lev=1&slev=5&obj=8&cmp=26&sreg=26&creg=2622000> [Accessed 2025-05-25]
65. Mansournia MA, Jewell NP, Greenland S. Case-control matching: effects, misconceptions, and recommendations. *Eur J Epidemiol.* Jan 2018;33(1):5-14. [doi: [10.1007/s10654-017-0325-0](https://doi.org/10.1007/s10654-017-0325-0)] [Medline: [29101596](https://pubmed.ncbi.nlm.nih.gov/29101596/)]
66. Stürmer T, Brenner H. Degree of matching and gain in power and efficiency in case-control studies. *Epidemiology (Sunnyvale).* Jan 2001;12(1):101-108. [doi: [10.1097/00001648-200101000-00017](https://doi.org/10.1097/00001648-200101000-00017)] [Medline: [11138803](https://pubmed.ncbi.nlm.nih.gov/11138803/)]
67. VanderWeele TJ, Robins JM. Four types of effect modification: a classification based on directed acyclic graphs. *Epidemiology (Sunnyvale).* Sep 2007;18(5):561-568. [doi: [10.1097/EDE.0b013e318127181b](https://doi.org/10.1097/EDE.0b013e318127181b)] [Medline: [17700242](https://pubmed.ncbi.nlm.nih.gov/17700242/)]
68. Weinberg CR. Can DAGs clarify effect modification? *Epidemiology (Sunnyvale).* Sep 2007;18(5):569-572. [doi: [10.1097/EDE.0b013e318126c11d](https://doi.org/10.1097/EDE.0b013e318126c11d)] [Medline: [17700243](https://pubmed.ncbi.nlm.nih.gov/17700243/)]
69. Williams TC, Bach CC, Matthiesen NB, Henriksen TB, Gagliardi L. Directed acyclic graphs: a tool for causal studies in paediatrics. *Pediatr Res.* Oct 2018;84(4):487-493. [doi: [10.1038/s41390-018-0071-3](https://doi.org/10.1038/s41390-018-0071-3)] [Medline: [29967527](https://pubmed.ncbi.nlm.nih.gov/29967527/)]
70. Evans D, Chaix B, Lobbedez T, Verger C, Flahault A. Combining directed acyclic graphs and the change-in-estimate procedure as a novel approach to adjustment-variable selection in epidemiology. *BMC Med Res Methodol.* Oct 11, 2012;12(1):156. [doi: [10.1186/1471-2288-12-156](https://doi.org/10.1186/1471-2288-12-156)] [Medline: [23058038](https://pubmed.ncbi.nlm.nih.gov/23058038/)]
71. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. *Ann Intern Med.* Aug 15, 2017;167(4):268-274. [doi: [10.7326/M16-2607](https://doi.org/10.7326/M16-2607)] [Medline: [28693043](https://pubmed.ncbi.nlm.nih.gov/28693043/)]
72. Mathur MB, Ding P, Riddell CA, VanderWeele TJ. Web site and R package for computing e-values. *Epidemiology (Sunnyvale).* Sep 2018;29(5):e45-e47. [doi: [10.1097/EDE.0000000000000864](https://doi.org/10.1097/EDE.0000000000000864)] [Medline: [29912013](https://pubmed.ncbi.nlm.nih.gov/29912013/)]
73. Wei R, Wang J, Su M, et al. Missing value imputation approach for mass spectrometry-based metabolomics data. *Sci Rep.* Jan 12, 2018;8(1):663. [doi: [10.1038/s41598-017-19120-0](https://doi.org/10.1038/s41598-017-19120-0)] [Medline: [29330539](https://pubmed.ncbi.nlm.nih.gov/29330539/)]
74. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B.* Jan 1, 1995;57(1):289-300. [doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)]
75. Zhu G, Wen Y, Cao K, He S, Wang T. A review of common statistical methods for dealing with multiple pollutant mixtures and multiple exposures. *Front Public Health.* 2024;12:1377685. [doi: [10.3389/fpubh.2024.1377685](https://doi.org/10.3389/fpubh.2024.1377685)] [Medline: [38784575](https://pubmed.ncbi.nlm.nih.gov/38784575/)]
76. Vuong AM, Xie C, Jandarov R, et al. Prenatal exposure to a mixture of persistent organic pollutants (POPs) and child reading skills at school age. *Int J Hyg Environ Health.* Jul 2020;228:113527. [doi: [10.1016/j.ijheh.2020.113527](https://doi.org/10.1016/j.ijheh.2020.113527)] [Medline: [32521479](https://pubmed.ncbi.nlm.nih.gov/32521479/)]
77. Schelldorfer J, Meier L, Bühlmann P. GLMMLasso: an algorithm for high-dimensional generalized linear mixed models using ℓ_1 -penalization. *J Comput Graph Stat.* Apr 3, 2014;23(2):460-477. [doi: [10.1080/10618600.2013.773239](https://doi.org/10.1080/10618600.2013.773239)]
78. Schubert B, Marks D, Sander C. GroupGELMMnet - generalized network-based elastic-net linear mixed model. GitHub. 2017. URL: <https://github.com/debbiemarkslab/GELMMnet> [Accessed 2025-05-25]
79. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol.* Dec 15, 2010;172(12):1339-1348. [doi: [10.1093/aje/kwq332](https://doi.org/10.1093/aje/kwq332)] [Medline: [21036955](https://pubmed.ncbi.nlm.nih.gov/21036955/)]
80. Aung MT, Song Y, Ferguson KK, et al. Application of an analytical framework for multivariate mediation analysis of environmental data. *Nat Commun.* Nov 6, 2020;11(1):5624. [doi: [10.1038/s41467-020-19335-2](https://doi.org/10.1038/s41467-020-19335-2)] [Medline: [33159049](https://pubmed.ncbi.nlm.nih.gov/33159049/)]
81. Liu Y, Lam KF, Wu JT, Lam TT. Geographically weighted temporally correlated logistic regression model. *Sci Rep.* Jan 23, 2018;8(1):1417. [doi: [10.1038/s41598-018-19772-6](https://doi.org/10.1038/s41598-018-19772-6)] [Medline: [29362396](https://pubmed.ncbi.nlm.nih.gov/29362396/)]

82. Matthews SA, Yang TC. Mapping the results of local statistics: using geographically weighted regression. *Demogr Res.* Mar 2, 2012;26:151-166. [doi: [10.4054/DemRes.2012.26.6](https://doi.org/10.4054/DemRes.2012.26.6)] [Medline: [25578024](https://pubmed.ncbi.nlm.nih.gov/25578024/)]
83. Chin JY, Godwin C, Parker E, et al. Levels and sources of volatile organic compounds in homes of children with asthma. *Indoor Air.* Aug 2014;24(4):403-415. [doi: [10.1111/ina.12086](https://doi.org/10.1111/ina.12086)] [Medline: [24329990](https://pubmed.ncbi.nlm.nih.gov/24329990/)]

Abbreviations

BTEX: benzene, toluene, ethylbenzene, and xylene

CLEAR: Center for Leadership in Environmental Awareness and Research

EDEN: *Etude des Déterminants pré et post natalis précoces du développement psychomoteur et de la santé de l'ENfant*

EMR: electronic medical record

GA: gestational age

HFH: Henry Ford Health

HIPAA: Health Insurance Portability and Accountability Act

MERL: Molecular Epidemiology Research Laboratory

PTB: preterm birth

REDCap: Research Electronic Data Capture

VOC: volatile organic compound

WHS: Women's Health Services

Edited by Amy Schwartz; peer-reviewed by Vinay Pal; submitted 06.Nov.2025; final revised version received 28.Apr.2026; accepted 29.Apr.2026; published 24.Jun.2026

Please cite as:

Cassidy-Bushrow AE, Leydet E, Levin AM, Lu M, Straughen JK

Investigating the Association Between Volatile Organic Compounds and Preterm Birth in Detroit, Michigan: Protocol for the Center for Leadership in Environmental Awareness and Research Birth Cohort Study

JMIR Res Protoc 2026;15:e87272

URL: <https://www.researchprotocols.org/2026/1/e87272>

doi: [10.2196/87272](https://doi.org/10.2196/87272)

© Andrea E Cassidy-Bushrow, Emily Leydet, Albert M Levin, Mei Lu, Jennifer K Straughen. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 24.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.