

The spatial and contextual exposome and subtypes of hypertensive disorders of pregnancy: a double machine learning-based analysis

Hui Hu^{1,*}, Claire L. Leiser¹, Xing He^{2,3}, Jaime E. Hart^{4,5}, Francine Laden^{1,4,5}, Cui Tao⁶, Jiang Bian^{2,3}

¹Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, United States

²Department of Biostatistics and Health Data Science, Indiana University School of Medicine, Indianapolis, IN 46202, United States

³Regenstrief Institute, Indianapolis, IN 46202, United States

⁴Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA 02115, United States

⁵Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, United States

⁶Department of Artificial Intelligence and Informatics, Mayo Clinic, Jacksonville, FL 32224, United States

*Corresponding author: Hui Hu, Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Ave, Room 330, Boston, MA 02115, United States (hui.hu@channing.harvard.edu)

Abstract

Hypertensive disorders of pregnancy (HDP) are a leading cause of maternal and perinatal morbidity, yet modifiable environmental risk factors remain poorly characterized. Prior studies typically have only examined a limited number of exposures and have rarely distinguished HDP subtypes (ie, gestational hypertension, preeclampsia, eclampsia, and chronic hypertension with or without superimposed preeclampsia) or accounted for residential mobility during pregnancy. To address these gaps, we conducted a spatial and contextual exposome study of HDP using linked electronic health records (EHR) and vital statistics data in Florida. We analyzed 686 412 singleton pregnancies conceived between 2013 and 2018, using computable phenotyping to distinguish HDP subtypes. A total of 245 spatial and contextual exposome measures spanning natural, built, and social environments were linked to residential histories from conception through gestational week 19. Using a two-phase double machine learning (DML) framework with exposure-specific, directed acyclic graph-guided confounder adjustment, we conducted discovery and replication analyses, followed by multi-treatment DML to estimate effect sizes. In Phase 1, 26 exposome measures replicated for gestational hypertension and 34 for overall HDP. In Phase 2, 12 measures remained associated with gestational hypertension and 11 with overall HDP, including air toxicants, meteorological factors, ultraviolet radiation, neighborhood crime indicators, environmental noise, and proximity to coastline. No exposures passed multiple-comparison thresholds for preeclampsia or eclampsia. These findings demonstrate that the spatial and contextual exposome contributes to HDP in a subtype-specific manner. Integrating EHR-linked phenotyping, residential mobility, and causal machine-learning methods offers a scalable framework for identifying environmental factors relevant to HDP prevention.

Key words: exposome; spatial; causal machine learning; hypertensive disorders of pregnancy; gestational hypertension; preeclampsia; eclampsia.

Introduction

Hypertensive disorders of pregnancy (HDP), including gestational hypertension, preeclampsia, eclampsia, and chronic hypertension with or without superimposed preeclampsia, are among the most common medical complications of pregnancy and contribute substantially to maternal and perinatal morbidity and mortality worldwide.¹⁻⁸ Women with HDP face elevated risks of adverse pregnancy outcomes such as preterm birth,⁹ fetal growth restriction,¹⁰ and perinatal death,^{11,12} as well as increased long-term risk of cardiovascular disease (CVD) later in life.¹³⁻¹⁷ Given the burden of HDP in many populations and the limited number of effective preventive strategies, identification of modifiable risk factors beyond known clinical and behavioral characteristics is needed in order to inform primary prevention and risk stratification in clinical practice. Emerging evidence

points to an important role of broader spatial and contextual environmental factors in identifying risk of HDP, with large geographic disparities in HDP incidence suggesting that where women live may be a key determinant of disease risk.^{18,19}

A growing body of work has evaluated specific environmental factors in relation to HDP, including ambient air pollution,²⁰ heavy metals,²¹ temperatures,^{22,23} neighborhood deprivation,²⁴ crime,²⁵ walkability,²⁶ green space,²⁷ and food access.²⁸ These studies have provided important evidence that spatial and contextual factors contribute to HDP risk, but have generally focused on a limited number of preselected exposures independently, without considering the broader “exposome”—the totality of non-genetic exposures across the life course.²⁹ The spatial and contextual exposome, which encompasses features of the natural, built, and social environments that can be assessed from

Received: December 23, 2025; Revised: February 26, 2026; Accepted: February 28, 2026

© The Author(s) 2026. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

preexisting spatial and contextual data linked to individuals,^{30,31} offers a powerful framework for identifying novel environmental determinants of HDP at scale.

Our prior work has demonstrated the feasibility of conducting large-scale external exposome-wide association studies (ExWAS) for HDP using statewide birth records data,³² incorporating hundreds of spatially and temporally resolved exposure variables and agnostic screening strategies. However, our study and others investigating this question have treated HDP as a composite outcome, combining gestational hypertension and preeclampsia-eclampsia, and did not distinguish between specific HDP subtypes and severity. HDP with different subtypes and severity (eg, gestational hypertension versus severe preeclampsia) may reflect distinct underlying pathophysiological processes.^{33,34} For example, some studies suggest that certain environmental exposures, such as air pollution, may be more strongly associated with milder forms of preeclampsia than with severe disease.^{35,36} Aggregating all HDP into a single composite outcome may therefore obscure important subtype-specific associations and limit our ability to identify high-risk women and tailor prevention strategies. Yet, due in part to limitations of commonly used data sources such as birth records and administrative claims—which often lack detailed clinical information—most environmental studies have not been able to differentiate HDP subtypes and severity.

Another key methodological challenge is how to accurately characterize residential exposures during pregnancy. Most population-based studies rely on a single residential address at delivery to assign environmental exposures, implicitly assuming that women remain at the same location throughout pregnancy. However, several studies have shown that ignoring residential mobility during pregnancy can result in exposure measurement error and subsequently biased effect estimates.^{37–39} Leveraging data sources that include detailed residential histories, such as electronic health records (EHR), is therefore critical to improve exposure assessment and reduce measurement error in studies of environmental determinants of HDP.

Beyond exposure assessment, there are important analytic limitations in current exposome literature. ExWAS approaches typically rely on high-dimensional yet essentially conventional regression models, applied exposure-by-exposure, with a common set of covariates assumed to adequately control confounding for all exposures under study.³⁰ Adjusting for a single uniform confounder set across diverse environmental exposures may be problematic because different exposures arise from distinct causal pathways and may therefore have different confounding structures. For example, air pollution is strongly influenced by traffic density and meteorology,⁴⁰ whereas green space is largely shaped by neighborhood and land-use context.⁴¹ Although both exposures may relate to health outcomes, the upstream determinants and potential confounders differ. Directed acyclic graphs (DAGs) provide a formal framework to represent hypothesized causal relationships and identify minimally sufficient adjustment sets tailored to each exposure, thereby reducing the risk of residual confounding or over-adjustment.⁴² However, even with DAG-informed confounder selection, modeling the complex, high-dimensional, and correlated nature of exposome data remains challenging. Traditional causal approaches, including parametric g-computation and propensity score-based methods, rely on correct model specification and may not perform well in settings characterized by high dimensionality, multicollinearity, and nonlinear relationships.^{43–45} Double machine learning (DML) offers an attractive alternative: by combining flexible

machine-learning algorithms for nuisance prediction with orthogonalized estimation of exposure effects, DML can accommodate high-dimensional confounding, complex functional forms, and correlated exposures while preserving valid statistical inference under appropriate assumptions.^{46,47} Yet, to our knowledge, formal DML-based causal inference has not been applied in the context of the spatial and contextual exposome or used to investigate HDP outcomes.

In this study, we address the aforementioned limitations by leveraging a large statewide linked EHR-vital statistics birth and fetal death records dataset from Florida, with detailed clinical information from EHR to distinguish HDP subtypes. We then reconstructed residential histories during pregnancy to assign time-resolved exposures across multiple address records. We define confounder sets tailored to different classes of spatial and contextual factors using exposure-specific causal diagrams and apply a DML framework to estimate associations between these exposures and HDP outcomes while flexibly adjusting for high-dimensional covariates. By integrating rich environmental data, granular HDP phenotyping, residential mobility, and modern causal machine-learning methods, this DML-based analysis aims to provide a more nuanced understanding of how the spatial and contextual exposome contributes to the development of specific HDP subtypes, and to identify potential targets for prevention and policy interventions.

Methods

Study population

We carried out a retrospective cohort study using a statewide dataset in Florida, USA that linked EHR with vital statistics birth and fetal death records. Data sources included the 2013–2019 Florida Vital Statistics Birth Records (VSBR) and Vital Statistics Fetal Death Records (VSFDR), obtained from the Bureau of Vital Statistics at the Florida Department of Health (<http://www.floridahealth.gov/certificates/>), as well as EHR data from the OneFlorida+ Clinical Research Network.⁴⁸ Record linkage was completed through a privacy-preserving system (Datavant) based on hashed personal identifiers, including name, date of birth, sex, and 5-digit ZIP code. Because maternal last names may change over time, we performed separate linkage procedures using both maiden names and surnames recorded at delivery as recorded in the VSBR/VSFDR. Estimated conception date was calculated primarily from delivery date and gestational age; when gestational age was missing, we used the reported last menstrual period to estimate conception. The initial linked dataset included 802 955 pregnancies with conception dates between January 1, 2013, and December 31, 2018. We then excluded multifetal pregnancies ($n=24\ 688$), given their higher risk of HDP compared with singleton pregnancies,⁴⁹ and records with implausible gestational ages (<20 or >45 weeks; $n=344$), resulting in a sample of 777 923 pregnancies. Consistent with prior studies of environmental exposures and HDP, as we focused on exposure to spatial and contextual factors from conception through the end of gestational week 19, we excluded pregnancies with chronic hypertension (with or without superimposed preeclampsia, $n=86\ 362$) and those classified as unspecified HDP ($n=5149$). A total of 686 412 pregnancies were included in the analyses.

Assessment of HDP subtypes

We identified hypertensive disorders of pregnancy (HDP) subtypes using data from both the EHR and the VSBR/VSFDR, drawing on diagnostic codes, blood pressure measurements,

antihypertensive medication records, and relevant laboratory findings. The HDP subtypes examined included chronic hypertension, gestational hypertension, mild and severe preeclampsia, eclampsia, superimposed preeclampsia on chronic hypertension, and unspecified HDP. The ICD-9-CM and ICD-10-CM codes used to define these conditions in the EHR are listed in [Table S1](#). Proteinuria was ascertained from urine and serum laboratory results obtained between 20 weeks' gestation and delivery, using LOINC codes and threshold values summarized in [Table S2](#). Use of antihypertensive medications was identified through RxNorm Concept Unique Identifiers (RxCUIs) established in prior work.⁵⁰ The final rule-based phenotyping algorithm, shown in [Table S3](#), was adapted from a recently developed approach that integrates information from both EHR and VSB/VSFDR sources.⁵¹

Assessment of the spatial and contextual exposome

To assess the spatial and contextual exposome for each pregnancy from conception through the end of gestational week 19, we assembled a broad set of environmental factors characterizing the natural, built, and social environment from multiple publicly available data sources ([Table 1](#)). Residential histories were derived from EHR address history records captured during clinical encounters. Although address records include start and end dates, these dates may not always represent true move-in and move-out dates (eg, when the move-in date is unknown, the EHR defaults the start date to the record creation date in the source system). We then spatiotemporally linked geocoded residential histories to exposome measures using circular buffers of 270-m (for all exposome measures) and 1230-m (additionally for green space, noise, and light at night) around each address. The 270-m buffer approximates the immediate residential environment, while the 1230-m buffer represents a walkable neighborhood area and was additionally applied to green space, noise, and nighttime light to capture accessible environmental context.⁵² Exposures were summarized within each buffer and combined across residential locations using area- and time-weighted averages over the spatiotemporal exposure window. [Table 1](#) shows a summary of the spatial and contextual exposome data sources. A total of 315 spatial and contextual exposome measures from 16 data sources covering 12 exposure classes were generated.

Natural environment

We obtained daily census tract-level ambient concentrations of fine particulate matter (PM_{2.5}) and ozone (O₃) from the U.S. Environmental Protection Agency (USEPA) Fused Air Quality Surface Using Downscaling (FAQSD) files.⁵³ These estimates were generated using a Bayesian space-time downscaler that integrates 12-km gridded estimates from the Models-3/Community Multiscale Air Quality (CMAQ) model with daily monitoring data for O₃ and PM_{2.5} from the National Air Monitoring Stations and State and Local Air Monitoring Stations (NAMS/SLAMS).⁵⁴ To characterize PM_{2.5} chemical composition, we included biweekly estimates of key PM_{2.5} constituents—sulfate (SO₄²⁻), ammonium (NH₄⁺), nitrate (NO₃⁻), organic matter (OM), black carbon (BC), mineral dust (DUST), and sea-salt (SS)—as well as the contribution from biomass burning. These measures were obtained from the Atmospheric Composition Analysis Group (Washington University in St Louis) at a 0.01° longitude/latitude spatial resolution for 2013–2019.⁵⁵ Measures of air toxicants were derived from AirToxScreen, previously known as the National Air Toxics Assessment (NATA), which was developed using a national emissions inventory of outdoor sources of air toxics.⁵⁶ Census

tract-level exposure estimates of 175 air toxicants available for 2014, 2018, and 2019 were obtained, and log-linear interpolations were performed to derive estimates in 2013 and 2015–2017. Daily meteorological measures were obtained from the PRISM Climate Group (Oregon State University) at 800-m spatial resolution for 2013–2019.⁵⁷ Daily heat index was then derived using daily mean temperature and relative humidity.⁵⁸ Measures on ultraviolet radiation were obtained from the Tropospheric Emission Monitoring Internet Service (TEMIS; European Space Agency) at 0.25° spatial resolution during 2013–2019. Proximity to blue space was quantified using the National Hydrography Dataset Plus Version 2 (U.S. Geological Survey).⁵⁹ Proximity to five different surface water features (ie, flowlines, waterbodies, areal or costal water, coastline, or any surface water features) were calculated.

Built environment

We assessed green space using the Normalized Difference Vegetation Index (NDVI), which was derived from MODIS/Terra and MODIS/Aqua satellite products at 250-m resolution.⁶⁰ The satellite data were processed by mosaicking tiles to a common grid covering Florida and applying quality and pixel reliability flags. Terra and Aqua time series were bias-corrected, blended to an effective 8-day temporal resolution, and short gaps (≤32 days) were linearly interpolated. Walkability was assessed using the USEPA National Walkability Index at the Census block group level.⁶¹ This index is a cross-sectional measure ranging from 1 to 20, with higher values reflecting greater walkability. We obtained tract-level indicators of food access from the US Department of Agriculture (USDA) Food Access Research Atlas for 2010, 2015, and 2019,⁶² and linear interpolation was performed to construct measures in gap years. Vacant land was characterized using the aggregated U.S. Postal Service address vacancy data from the US Department of Housing and Urban Development (USHUD), which provide tract-level vacancy indicators at a quarterly (3-month) temporal resolution.⁶³ These measures capture residential and business address vacancies and have been used as proxies for neighborhood disinvestment and physical disorder. We assessed road proximity using TIGER/Line road network data from the U.S. Census Bureau. Distances from addresses to four categories of major road classes were calculated, including A1 roads (ie, primary highways with limited access), A2 roads (ie, primary roads without limited access), A3 roads (ie, secondary and connecting roads), and any of the A1, A2, or A3 roads. Noise exposure was assessed using ambient sound pressure levels derived by the National Park Service Natural Sounds and Night Skies Division, which provide modeled sound-level estimates at 270-m resolution,^{64,65} including three A-weighted median (L50) sound pressure level metrics: existing conditions (representing the sound level from all natural and anthropogenic sources), natural conditions (representing the level expected in the absence of human-generated noise), and impact conditions (the difference between existing and natural conditions, capturing the incremental contribution of anthropogenic noise). Light at night was quantified using annual nighttime lights products from the Visible Infrared Imaging Radiometer Suite (VIIRS) Version 2 (NASA) at approximately 15-second spatial resolution.⁶⁶

Social environment

Neighborhood socioeconomic status (SES) was assessed using the Neighborhood Deprivation Index (NDI),⁶⁷ a validated measure derived from American Community Survey 5-year estimates at the Census block group level for 2011–2021. The NDI integrates multiple indicators of socioeconomic conditions, including poverty,

Table 1. Summary of spatial and contextual exposome measures.

Category	Data source	Time period	Spatial scale	Temporal scale	Number of measures	Number of variables ^a
PM _{2.5} and O ₃	Fused Air Quality Surface Using Downscaling Files, USEPA	2013-2019	Census tract	1-day	2	2
PM _{2.5} compositions	Atmospheric Composition Analysis Group, WUSTL	2013-2019	0.01 degree in lon/lat	2-week	24	18
Air toxicants	National Air Toxic Assessment/Air Toxics Screening, USEPA	2014, 2018, 2019	Census tract	1-year	175	119
Meteorology	PRISM, Oregon State University	2013-2019	800m	1-day	9	8
Ultraviolet	Tropospheric Emission Monitoring Internet Service, ESA	2013-2019	0.25 deg in lon/lat	1-day	4	1
Blue space	National Hydrography Dataset Plus version 2, USGS	CS	Vector	CS	5	5
Green space	MODIS/Terra and MODIS/Aqua Normalized Difference Vegetation Index, NASA	2013-2019	250m	16-day	2	2
Walkability	Walkability Index, USEPA	2019	Census block group	Cross-sectional	1	1
Food access	Food Access Research Atlas, USDA	2010, 2015, 2019	Census tract	1-year	44	42
Noise	Natural Sounds and Night Skies Division, US NPS	CS	270m	CS	6	5
Land vacancy	Aggregated USPS Administrative Data on Address Vacancies, USHUD	2013-2019	Census tract	3-month	19	18
Road proximity	TIGERLine, US Census Bureau	2013-2019	Vector	1-year	3	3
Light at night	Visible Infrared Imaging Radiometer Suite version 2, NASA	2013-2019	15 second in lon/lat	1-year	2	2
Neighborhood deprivation	American Community Survey, US Census Bureau	2011-2021	Census block group	5-year	1	1
Social capital	Census Business Pattern, US Census Bureau	2013-2019	ZCTA5	1-year	10	10
Crime and safety	Uniform Crime Reporting Program, FBI	2013-2019	County	1-year	8	8

^aNumber of variables after removing 49 measures with number of unique values < 0.1% of the total sample size and 21 measures with absolute correlations > 0.99 with another measure.

Abbreviations: CS: Cross-sectional; ESA, European Space Agency; FBI, Federal Bureau of Investigation; MODIS, The Moderate Resolution Imaging Spectroradiometer; NASA, The National Aeronautics and Space Administration; NPS, National Park Service; USDA, United States Department of Agriculture; USEPA, United States Environmental Protection Agency; USGS, United States Geological Survey; USHUD, United States Department of Housing and Urban Development; WUSTL, Washington University at St Louis; ZCTA5, 5-digit ZIP Code Tabulation Areas.

occupation, housing, employment, education, racial composition, and residential stability, into a single standardized score, with higher values indicating more deprived neighborhoods. For each ACS 5-year file, we used the midpoint year of the reporting period as the index year (eg, 2013-2017 ACS assigned to 2015) when linking NDI to pregnancies. Social capital was assessed using contextual measures constructed from the Census Business Patterns data at the 5-digit ZIP Code Tabulation Area (ZCTA) level for 2013-2019. We used North American Industry Classification System (NAICS) codes to derive densities of organizations such as religious, civic, recreational, business, and professional groups, and aggregated them into ten social capital indicators.⁶⁸ Crime and safety were represented by eight annual county-level crime measures from the Federal Bureau of Investigation Uniform Crime Reporting Program (2013-2019), including overall crime and specific categories such as violent and property crimes.⁶⁹

Covariates

We considered a range of covariates derived from both the EHR and VSBP/VSFDR data, including maternal age at delivery (<20, 20-24, 25-29, 30-34, 35-39, and ≥40 years old), race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian/Pacific Islander, and others), education level (< high school, high school or equivalent, and >high school), marital status at delivery (yes/no), pre-pregnancy BMI (underweight: <18.5 kg/m³, normal: 18.5-24.9 kg/m³, overweight: 25-29.9 kg/m³, obese: ≥30 kg/m³),

smoking during pregnancy (yes/no), timing of prenatal care initiation (no care, first trimester, second trimester, third trimester, or yes with unknown start date), insurance type (Medicaid, private insurance, self-pay, and others), participation in the US Department of Agriculture's Women, Infants, and Children program (WIC; yes/no), pre-pregnancy depression (yes/no), parity (nulliparous/non-nulliparous), and month/year of conception. Number of residential moves (determined from residential histories) and urbanicity (determined by linking residential histories to 2010 Census urban area classifications, with urbanized areas and urban clusters classified as urban and all other areas classified as rural) from conception through gestational week 19 were further considered in sensitivity analyses.

Statistical analyses

Maternal sociodemographic, behavioral, and clinical characteristics by HDP subtype were described. All continuous spatial and contextual exposome measures were standardized (ie, mean = 0 and standard deviation = 1). We further removed (1) 49 spatial and contextual exposome measures with very low variability (defined as having unique values representing <0.1% of the sample size, Table S4) because such sparse variation provides minimal exposure contrast and can yield unstable estimates and reduced statistical power in high-dimensional models, and (2) 21 measures with absolute correlations > 0.99 with another measure (Table S5). A total of 245 exposome measures were included in

the analyses. Missing values in both exposures and covariates were addressed using multiple imputation by chained equations implemented in the *mice* package in R.⁷⁰ Variables were included as predictors in the imputation models if, among individuals with missing values in the target variable, more than 40% had observed data on that predictor and it showed at least moderate association ($|r| > 0.4$) with either the variable being imputed or the indicator of missingness. Because the overall proportion of missing data was low and the cohort was large, we generated a single imputed dataset for the main analyses. We then used a two-phase DML framework to examine the relationships between the spatial and contextual exposome and HDP (overall and by subtype). To guide the selection of confounders for each exposome measure, we specified a directed acyclic graph (DAG, see [Figure S1](#) and [Table S6](#)) including all covariates and 12 exposure classes (ie, air pollution, climate, UV radiation, blue space, green space, walkability, food access, noise, land vacancy, road proximity, light at night, neighborhood deprivation, social capital, and crime and safety). Initial DAGs were independently drafted by two investigators based on prior literature and subject-matter knowledge, with discrepancies resolved through discussion and, when needed, consultation with a third investigator to achieve consensus. For each exposure class, we derived a minimal sufficient adjustment set using the *dagitty* package in R and then expanded these nodes to concrete variables (baseline covariates and exposures from other classes) to serve as potential confounders.⁷¹ [Figure 1](#) shows the flow chart summarizing the analysis pipeline.

In Phase 1, we split the analytic sample by year of conception into a discovery set (2013–2015 conceptions) and an independent replication set (2016–2018 conceptions) to provide temporally independent validation and to evaluate robustness across potential secular changes in environmental exposures and population characteristics. We implemented single-exposure DML using a logistic partially linear model for each exposure and each HDP outcome (individual HDP subtypes versus normotensive pregnancies, and overall HDP versus normotensive pregnancies).⁷² For each exposome measure, we modeled individual HDP subtypes and overall HDP as the outcomes separately, and the DAG-selected variables (specific for each exposome measure) as high-dimensional confounders. Following the logistic partially linear DML framework,⁷² we combined: (1) a machine-learned model for the exposure given confounders, and (2) a machine-learned model for the outcome probability given confounders, into an orthogonal score to estimate an exposure log-odds ratio that is robust to small errors in nuisance estimation. Both nuisance functions were fitted with gradient boosting (XGBoost)⁷³ using separate hyperparameter grids for the exposure regression and outcome classification tuned with 3-fold cross-validation. We used 5-fold cross-fitting to obtain out-of-fold nuisance predictions, and solved the orthogonal estimating equation for each exposure-outcome pair. Standard errors, 95% confidence intervals (CIs), and P-values were obtained using a multiplier bootstrap. Phase-1 screening was performed separately in the discovery and replication samples; we applied the Benjamini-Hochberg procedure to control the false discovery rate (FDR) across all exposure-outcome tests within each sample, and carried forward exposures that had FDR-adjusted P-values < 0.05 in both discovery and replication and showed consistent directions of association.

In Phase 2, we used a multi-treatment DML extension to jointly estimate associations for the subset of replicated exposures from Phase 1.⁷⁴ For each outcome, we constructed a multi-treatment design in which the selected standardized exposures

entered as a vector of “treatments,” and a single DAG-based confounder set was used to build the nuisance models. Similar to Phase 1, we then fit: (1) machine-learning models for all exposures given confounders, and (2) a classification model for HDP given confounders with XGBoost⁷³ and separate hyperparameter grids tuned with 3-fold cross-validation. 5-fold cross-fitting was used to obtain out-of-fold nuisance predictions. The final effect estimates were obtained by fitting a logistic regression for the outcome including all residualized exposures simultaneously, with the logit of the nuisance propensity as an offset, yielding mutually adjusted odds ratios (ORs) and 95% CIs for each exposure in the multi-exposure setting. Sensitivity analyses were conducted to assess the robustness of our findings by additionally adjusting for potential confounders, including the number of residential moves and urbanicity. All statistical analyses were conducted in R version 4.4.2 (RStudio Team 2020). This study received approval from the institutional review boards at the Mass General Brigham (2021P002672), the Florida Department of Health (2019-106), and the University of Florida (IRB201902383).

Results

[Table 2](#) presents maternal characteristics for the overall study population and according to HDP subtype. Among 686 412 singleton pregnancies with conception dates between 2013 and 2018 in the linked Florida EHR-vital statistics birth and fetal death records dataset, 66 095 HDP cases were identified (9.6%). Of these, 41 865 were gestational hypertension (6.1%), 10 566 were mild preeclampsia (1.5%), 11 811 were severe preeclampsia (1.7%), and 1853 were eclampsia (0.3%). Compared with pregnancies without HDP, pregnancies affected by HDP were more likely to occur among women aged < 20 years (9.6% versus 7.4%) or ≥ 40 years (2.8% versus 2.3%), and among non-Hispanic Black women (26.0% versus 21.1%), while Hispanic women comprised a smaller proportion of HDP cases (30.2% versus 36.7%). Women with HDP were also more likely to have pre-pregnancy obesity (34.6% versus 21.8%) and to be covered by Medicaid (71.7% versus 69.8%), and they were more often nulliparous (52.5% versus 37.7%) and had a higher prevalence of pre-pregnancy depression (3.0% versus 2.6%). Across HDP subtypes, severe preeclampsia and eclampsia showed the highest proportions of non-Hispanic Black women (31.1% and 35.5%, respectively) and were less likely to have prenatal care initiated in the first trimester (46.2% and 49.9%, respectively); eclampsia had the highest prevalence of pre-pregnancy depression (7.3%). [Table S7](#) shows the distribution of spatial and contextual exposome measures included in the analyses, along with the percentage of missing values for each measure.

Volcano plots shown in [Figures 2](#) and [3](#) summarize the Phase 1 results based on the discovery and replication sets, respectively. [Tables S8–S12](#) present the estimated effect sizes for HDP subtypes and overall HDP in Phase 1. As shown in [Table S8](#), 91 and 65 exposome measures were significantly associated with gestational hypertension in the discovery and replication sets, respectively. After accounting for multiple comparisons, 26 measures that remained significant in both sets were retained in Phase 2. Similarly, 34 exposome measures were significantly associated with overall HDP in both the discovery and replication sets after multiple-comparison adjustment, as shown in [Table S12](#). Although no exposome measures remained significant after multiple-comparison adjustment for mild preeclampsia ([Table S9](#)), severe preeclampsia ([Table S10](#)), or eclampsia ([Table S11](#)), three measures were significant for severe preeclampsia in both

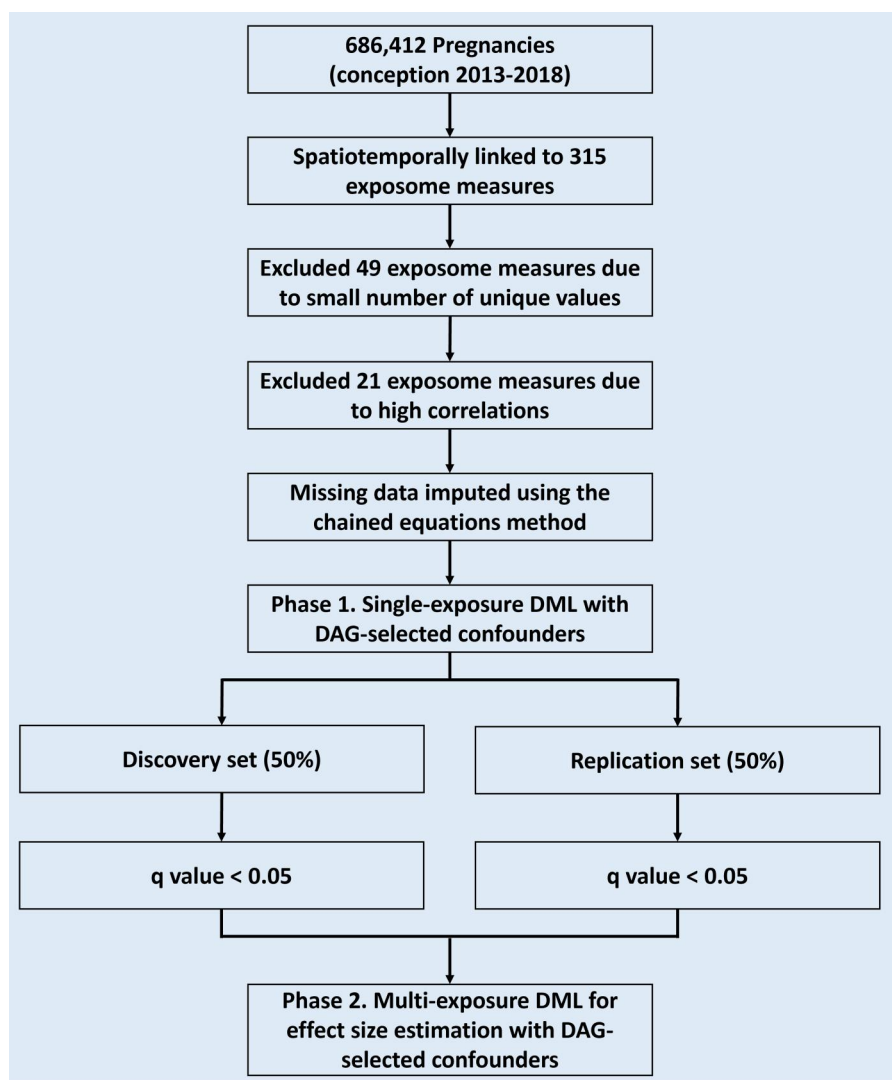


Figure 1. Flowchart of the analysis pipeline.

sets ($P < 0.05$) (Table S10), including two measures negatively associated with severe preeclampsia (ie, minimum relative humidity, natural A-weighted L50 sound pressure level), and arsenic compounds (inorganic including arsine), which are positively associated with severe preeclampsia.

In Phase 2, the exposome measures retained from Phase 1 were simultaneously included in a multi-treatment DML model after adjusting for confounders selected based on the DAG. Table 3 presents the adjusted odds ratios (ORs) for gestational hypertension per standard deviation increase in each exposome measure, along with 95% confidence intervals (CIs). Among the 26 measures included in Phase 2, 12 remained significantly associated with gestational hypertension, spanning air toxicants ($n=5$), blue space ($n=1$), crime and safety ($n=2$), meteorology ($n=3$), and ultraviolet radiation ($n=1$). Specifically, significant associations were observed for hydrogen fluoride (hydrofluoric acid) (OR: 0.99, 95% CI: 0.98, 1.00), methyl bromide (bromomethane) (OR: 0.98, 95% CI: 0.98, 0.99), methyl isobutyl ketone (hexone) (OR: 0.97, 95% CI: 0.95, 1.00), 2,2,4-trimethylpentane (OR: 1.05, 95% CI: 1.03, 1.07), and catechol (OR: 1.03, 95% CI: 1.01, 1.05). The blue space measure proximity to coastline was also positively associated with gestational hypertension (OR: 1.04, 95% CI: 1.03, 1.05). For crime and safety, both burglary rate (OR:

1.02, 95% CI: 1.00, 1.04) and forcible sex offenses rate (OR: 1.01, 95% CI: 1.00, 1.02) were significantly associated with higher odds of gestational hypertension. Meteorological measures showing significant associations included maximum relative humidity (OR: 1.02, 95% CI: 1.01, 1.04), minimum relative humidity (OR: 0.95, 95% CI: 0.93, 0.98), and maximum temperature (OR: 0.92, 95% CI: 0.86, 1.00). Finally, a higher erythemal UV index was associated with lower odds of gestational hypertension (OR: 0.94, 95% CI: 0.90, 0.98).

Table 4 presents the Phase 2 results for overall HDP. Among the 34 measures included, 11 remained significantly associated with overall HDP in the multi-treatment DML model, including 7 air toxicants, 1 blue space measure, 1 meteorology measure, 1 noise measure, and 1 ultraviolet measure. Specifically, overall HDP was positively associated with the air toxicants selenium compounds (OR: 1.04, 95% CI: 1.03, 1.05), 2,2,4-trimethylpentane (OR: 1.04, 95% CI: 1.02, 1.06), and glycol ethers (OR: 1.03, 95% CI: 1.00, 1.05), as well as the blue space measure proximity to coastline (OR: 1.02, 95% CI: 1.02, 1.03) and the meteorology measure maximum relative humidity (OR: 1.02, 95% CI: 1.00, 1.03). Inverse associations were observed for the air toxicants hydrogen fluoride (hydrofluoric acid) (OR: 0.99, 95% CI: 0.98, 0.99), methyl bromide (bromomethane) (OR: 0.99, 95% CI: 0.98, 0.99), xylenes

Table 2. Maternal characteristics by overall HDP and HDP subtypes among singleton pregnancies with a conception date 2013-2018 in the EHR-vital statistics birth and fetal death records data in Florida (n = 686 412).

Characteristics	Total (n = 686 412)	No HDP (n = 620 317)	Overall HDP (n = 66 095)	Gestational hypertension (n = 41 865)	Mild preeclampsia (n = 10 566)	Severe preeclampsia (n = 11 811)	Eclampsia (n = 1853)
Age							
<20	52 436 (7.6)	46 067 (7.4)	6369 (9.6)	3619 (8.6)	1175 (11.1)	1341 (11.4)	234 (12.6)
20-24	189 340 (27.6)	170 754 (27.5)	18 586 (28.1)	11 344 (27.1)	3350 (31.7)	3365 (28.5)	527 (28.4)
25-29	207 324 (30.2)	188 648 (30.4)	18 676 (28.3)	12 221 (29.2)	2966 (28.1)	2988 (25.3)	501 (27.0)
30-34	150 287 (21.9)	136 609 (22.0)	13 678 (20.7)	9074 (21.7)	1933 (18.3)	2334 (19.8)	337 (18.2)
35-39	70 629 (10.3)	63 717 (10.3)	6912 (10.5)	4441 (10.6)	890 (8.4)	1382 (11.7)	199 (10.7)
≥40	16396 (2.4)	14522 (2.3)	1 874 (2.8)	1166 (2.8)	252 (2.4)	401 (3.4)	55 (3.0)
Race/ethnicity							
Non-Hispanic White	255 252 (37.2)	229 121 (36.9)	26 131 (39.5)	17 932 (42.8)	4052 (38.3)	3579 (30.3)	568 (30.7)
Non-Hispanic Black	148 200 (21.6)	130 996 (21.1)	17 204 (26.0)	10 259 (24.5)	2613 (24.7)	3674 (31.1)	658 (35.5)
Hispanic	247 905 (36.1)	227 932 (36.7)	19 973 (30.2)	11 828 (28.3)	3503 (33.2)	4073 (34.5)	569 (30.7)
Asian/Pacific Islander	14 689 (2.1)	13 669 (2.2)	1020 (1.5)	695 (1.7)	123 (1.2)	187 (1.6)	15 (0.8)
Other	15 613 (2.3)	14 196 (2.3)	1417 (2.1)	959 (2.3)	204 (1.9)	227 (1.9)	27 (1.5)
Missing	4753 (0.7)	4403 (0.7)	350 (0.5)	192 (0.5)	71 (0.7)	71 (0.6)	16 (0.9)
Education							
< High school	103 864 (15.1)	94 132 (15.2)	9732 (14.7)	5705 (13.6)	1686 (16.0)	1993 (16.9)	348 (18.8)
High school or equivalent	267 378 (39.0)	241 201 (38.9)	26 177 (39.6)	16 052 (38.3)	4519 (42.8)	4808 (40.7)	798 (43.1)
> High school	309 119 (45.0)	279 475 (45.1)	29 644 (44.9)	19 801 (47.3)	4266 (40.4)	4896 (41.5)	681 (36.8)
Missing	6051 (0.9)	5509 (0.9)	542 (0.8)	307 (0.7)	95 (0.9)	114 (1.0)	26 (1.4)
Marital Status							
Not married	422 884 (61.6)	380 535 (61.3)	42 349 (64.1)	25 971 (62.0)	7323 (69.3)	7704 (65.2)	1351 (72.9)
Married	262 052 (38.2)	238 459 (38.4)	23 593 (35.7)	15 828 (37.8)	3236 (30.6)	4040 (34.2)	489 (26.4)
Missing	1476 (0.2)	1323 (0.2)	153 (0.2)	66 (0.2)	7 (0.1)	67 (0.6)	13 (0.7)
Pre-pregnancy BMI							
Underweight (<18.5)	30 391 (4.4)	28 400 (4.6)	1991 (3.0)	1193 (2.8)	308 (2.9)	431 (3.6)	59 (3.2)
Normal (18.5-24.9)	283 276 (41.3)	262 955 (42.4)	20 321 (30.7)	12 730 (30.4)	3222 (30.5)	3832 (32.4)	537 (29.0)
Overweight (25.0-29.9)	174 038 (25.4)	157 626 (25.4)	16 412 (24.8)	10 488 (25.1)	2685 (25.4)	2836 (24.0)	403 (21.7)
Obese (≥30.0)	158 288 (23.1)	135 396 (21.8)	22 892 (34.6)	14 655 (35.0)	3782 (35.8)	3774 (32.0)	681 (36.8)
Missing	40 419 (5.9)	35 940 (5.8)	4479 (6.8)	2799 (6.7)	569 (5.4)	938 (7.9)	173 (9.3)
Smoking during pregnancy							
No	636 272 (92.7)	574 702 (92.6)	61 570 (93.2)	38 877 (92.9)	9810 (92.8)	11 169 (94.6)	1714 (92.5)
Yes	45 884 (6.7)	41 807 (6.7)	4077 (6.2)	2705 (6.5)	702 (6.6)	546 (4.6)	124 (6.7)
Missing	4256 (0.6)	3808 (0.6)	448 (0.7)	283 (0.7)	54 (0.5)	96 (0.8)	15 (0.8)
Prenatal care began							
No care	12 741 (1.9)	11 513 (1.9)	1228 (1.9)	685 (1.6)	130 (1.2)	351 (3.0)	62 (3.3)
First trimester	358 442 (52.2)	324 927 (52.4)	33 515 (50.7)	21 708 (51.9)	5426 (51.4)	5456 (46.2)	925 (49.9)
Second trimester	141 230 (20.6)	127 853 (20.6)	13 377 (20.2)	8449 (20.2)	2167 (20.5)	2352 (19.9)	409 (22.1)
Third trimester	38 377 (5.6)	34 900 (5.6)	3477 (5.3)	2368 (5.7)	494 (4.7)	532 (4.5)	83 (4.5)
Yes, no clear information	132 270 (19.3)	118 090 (19.0)	14 180 (21.5)	8473 (20.2)	2318 (21.9)	3036 (25.7)	353 (19.1)
Missing	3352 (0.5)	3034 (0.5)	318 (0.5)	182 (0.4)	31 (0.3)	84 (0.7)	21 (1.1)
Insurance type							
Medicaid	480 167 (70.0)	432 794 (69.8)	47 373 (71.7)	29 083 (69.5)	8486 (80.3)	8256 (69.9)	1548 (83.5)
Private insurance	158 275 (23.1)	143 561 (23.1)	14 714 (22.3)	10 547 (25.2)	1570 (14.9)	2411 (20.4)	186 (10.0)
Self-pay	29 840 (4.3)	27 416 (4.4)	2424 (3.7)	1346 (3.2)	308 (2.9)	713 (6.0)	57 (3.1)
Others	11 798 (1.7)	10 775 (1.7)	1023 (1.5)	608 (1.5)	128 (1.2)	263 (2.2)	24 (1.3)
Missing	6332 (0.9)	5771 (0.9)	561 (0.8)	281 (0.7)	74 (0.7)	168 (1.4)	38 (2.1)
WIC							
No	244 114 (35.6)	221 212 (35.7)	22 902 (34.7)	15 432 (36.9)	2927 (27.7)	4009 (33.9)	534 (28.8)
Yes	436 400 (63.6)	393 736 (63.5)	42 664 (64.5)	26 124 (62.4)	7565 (71.6)	7690 (65.1)	1285 (69.3)
Missing	5898 (0.9)	5369 (0.9)	529 (0.8)	309 (0.7)	74 (0.7)	112 (0.9)	34 (1.8)
Pre-pregnancy depression	18 031 (2.6)	16 058 (2.6)	1973 (3.0)	1235 (2.9)	304 (2.9)	299 (2.5)	135 (7.3)
Parity							
Nulliparous	268 813 (39.2)	234 135 (37.7)	34 678 (52.5)	21 256 (50.8)	5617 (53.2)	6963 (59.0)	842 (45.4)
Non-nulliparous	415 524 (60.5)	384 359 (62.0)	31 165 (47.2)	20 451 (48.8)	4919 (46.6)	4791 (40.6)	1004 (54.2)
Missing	2075 (0.3)	1823 (0.3)	252 (0.4)	158 (0.4)	30 (0.3)	57 (0.5)	7 (0.4)
Number of residential moves (from concep- tion through gesta- tional week 19)							
0	635 402 (92.6)	574 987 (92.7)	60 415 (91.4)	38 261 (91.4)	9778 (92.5)	10 704 (90.6)	1672 (90.2)
1	48 941 (7.1)	43 490 (7.0)	5451 (8.2)	3441 (8.2)	767 (7.3)	1066 (9.0)	177 (9.6)
2	2066 (0.3)	1838 (0.3)	228 (0.3)	160 (0.4)	22 (0.2)	42 (0.4)	4 (0.2)
3	3 (0.0)	2 (0.0)	1 (0.0)	1 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

(continued)

Table 2. (continued)

Characteristics	Total (n = 686 412)	No HDP (n = 620 317)	Overall HDP (n = 66 095)	Gestational hypertension (n = 41 865)	Mild preeclampsia (n = 10 566)	Severe preeclampsia (n = 11 811)	Eclampsia (n = 1853)
Urbanicity (from conception through gestational week 19)							
Always urban	626 929 (91.3)	567 376 (91.5)	59 553 (90.1)	37 422 (89.4)	9582 (90.7)	10 883 (92.1)	1666 (89.9)
Always rural	55 057 (8.0)	49 030 (7.9)	6027 (9.1)	4079 (9.7)	924 (8.7)	851 (7.2)	173 (9.3)
Mix of urban and rural	4426 (0.6)	3911 (0.6)	515 (0.8)	362 (0.9)	61 (0.6)	78 (0.7)	14 (0.8)
Month of conception							
January	59 284 (8.6)	53 681 (8.7)	5603 (8.5)	3558 (8.5)	919 (8.7)	970 (8.2)	156 (8.4)
February	54 183 (7.9)	49 131 (7.9)	5052 (7.6)	3177 (7.6)	822 (7.8)	912 (7.7)	141 (7.6)
March	60 835 (8.9)	55 002 (8.9)	5833 (8.8)	3731 (8.9)	887 (8.4)	1032 (8.7)	183 (9.9)
April	57 432 (8.4)	51 936 (8.4)	5496 (8.3)	3505 (8.4)	908 (8.6)	940 (8.0)	143 (7.7)
May	58 814 (8.6)	53 146 (8.6)	5668 (8.6)	3487 (8.3)	967 (9.2)	1025 (8.7)	189 (10.2)
June	54 216 (7.9)	48 917 (7.9)	5299 (8.0)	3339 (8.0)	905 (8.6)	913 (7.7)	142 (7.7)
July	55 081 (8.0)	49 727 (8.0)	5354 (8.1)	3415 (8.2)	851 (8.1)	930 (7.9)	158 (8.5)
August	54 323 (7.9)	49 063 (7.9)	5260 (8.0)	3324 (7.9)	832 (7.9)	949 (8.0)	155 (8.4)
September	53 004 (7.7)	47 881 (7.7)	5123 (7.8)	3297 (7.9)	767 (7.3)	930 (7.9)	129 (7.0)
October	57 886 (8.4)	52 263 (8.4)	5623 (8.5)	3512 (8.4)	912 (8.6)	1063 (9.0)	136 (7.3)
November	59 099 (8.6)	53 278 (8.6)	5821 (8.8)	3701 (8.8)	913 (8.6)	1052 (8.9)	155 (8.4)
December	62 255 (9.1)	56 292 (9.1)	5963 (9.0)	3819 (9.1)	883 (8.4)	1095 (9.3)	166 (9.0)
Year of conception							
2013	114 463 (16.7)	104 213 (16.8)	10 250 (15.5)	5714 (13.6)	2615 (24.7)	1680 (14.2)	241 (13.0)
2014	115 343 (16.8)	104 952 (16.9)	10 391 (15.7)	6072 (14.5)	2398 (22.7)	1667 (14.1)	254 (13.7)
2015	114 298 (16.7)	103 836 (16.7)	10 462 (15.8)	6947 (16.6)	1380 (13.1)	1783 (15.1)	352 (19.0)
2016	116 975 (17.0)	105 653 (17.0)	11 322 (17.1)	7695 (18.4)	1304 (12.3)	2027 (17.2)	296 (16.0)
2017	114 343 (16.7)	102 748 (16.6)	11 595 (17.5)	7477 (17.9)	1470 (13.9)	2309 (19.5)	339 (18.3)
2018	110 990 (16.2)	98 915 (15.9)	12 075 (18.3)	7960 (19.0)	1399 (13.2)	2345 (19.9)	371 (20.0)

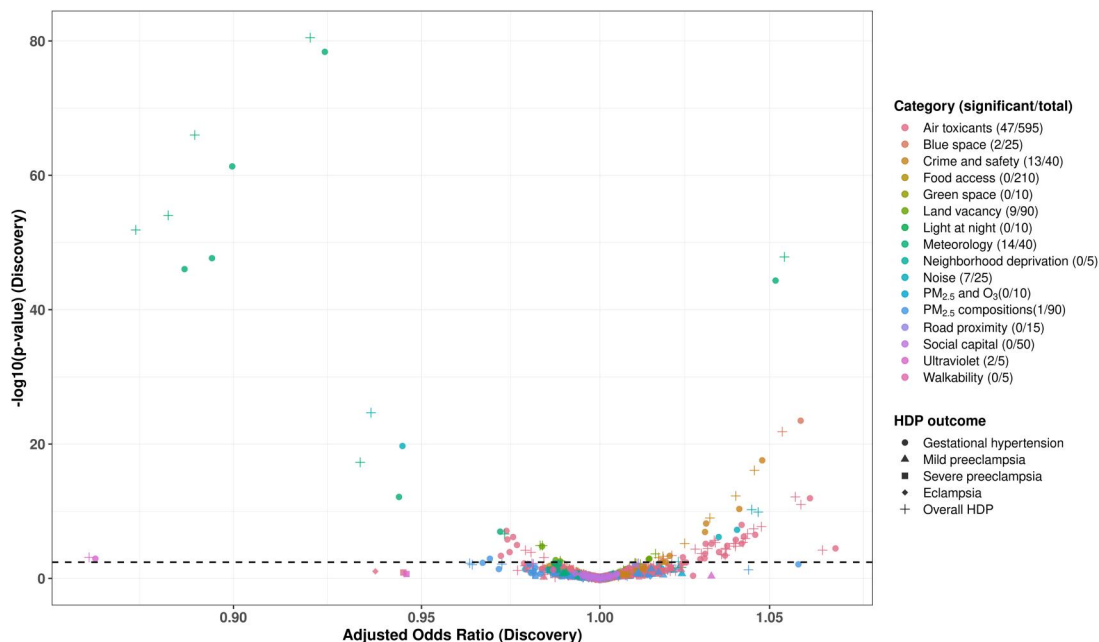


Figure 2. Volcano plot showing the results from Phase 1 analysis of the discovery set on individual HDP subtypes and overall HDP among singleton pregnancies with a conception date between 2013 and 2015 in the linked EHR-vital statistics birth and fetal death records in Florida ($n = 344\ 104$).

(mixed isomers) (OR: 0.95, 95% CI: 0.92, 0.98), and 1,1,1-trichloroethane (OR: 0.97, 95% CI: 0.95, 0.99), as well as the noise measure natural A-weighted L50 sound pressure level (OR: 0.97, 95% CI: 0.97, 0.99) and the ultraviolet measure erythral UV index (OR: 0.94, 95% CI: 0.91, 0.98).

Results from the sensitivity analyses (Tables S13 and S14) were largely consistent with the primary analyses after additional adjustment for the number of residential moves and urbanicity. Minor differences were observed: methyl isobutyl

ketone (primary P-value: 0.04; sensitivity P-value: 0.13) and maximum temperature (primary P-value: 0.04; sensitivity P-value: 0.22) were no longer statistically significant for gestational hypertension, and glycol ethers (primary P-value: 0.04; sensitivity P-value: 0.25) was no longer significant for overall HDP, although effect estimates remained similar. Conversely, several exposures that were not statistically significant in the primary analyses became significant in the sensitivity analyses, including hexane (primary P-value: 0.05; sensitivity P-value: 0.003), quinone

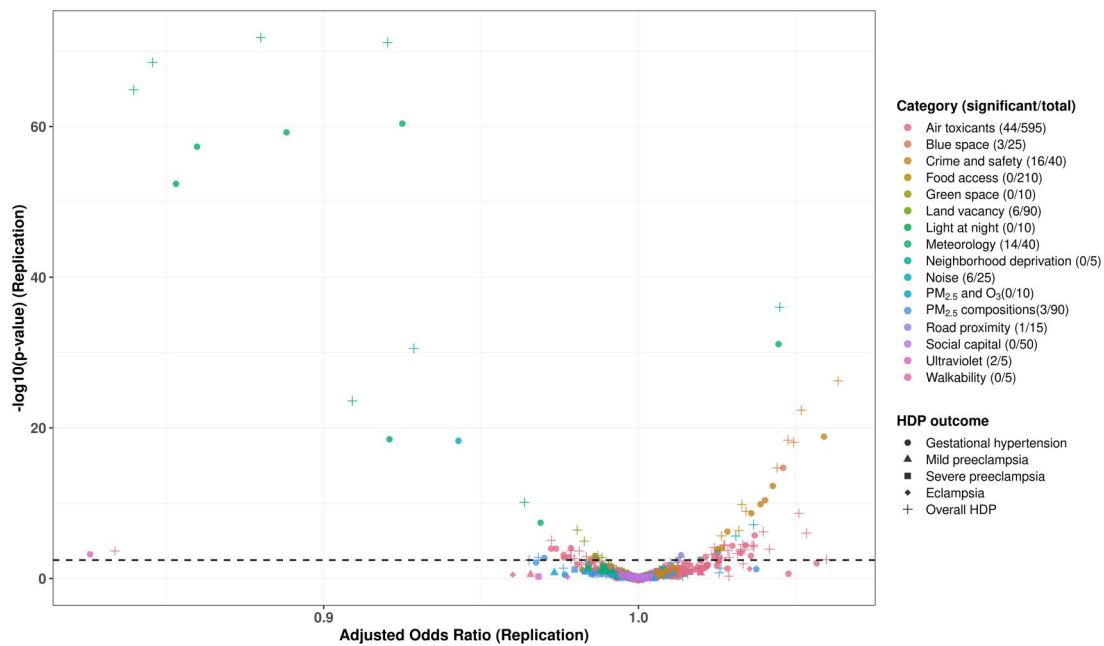


Figure 3. Volcano plot showing the results from Phase 1 analysis of the replication set on individual HDP subtypes and overall HDP among singleton pregnancies with a conception date between 2016 and 2018 in the linked EHR-vital statistics birth and fetal death records in Florida ($n = 342\ 308$).

(primary P -value: 0.33; sensitivity P -value: 0.02), 2,4-dinitrotoluene (primary P -value: 0.18; sensitivity P -value: 0.003), catechol (primary P -value: 0.06; sensitivity P -value: 0.02), forcible sex offense rate (primary P -value: 0.06; sensitivity P -value: 0.04) and murder rate (primary P -value: 0.14; sensitivity P -value: 0.01). Overall, these changes largely reflected shifts from marginal non-significance to statistical significance (or vice versa), while the magnitude and direction of the effect estimates remained stable.

Discussion

In this large Florida cohort with linked birth/fetal death records and EHR data, we found that multiple early-pregnancy environmental exposures were associated with HDP. Using a two-phase DML-based approach, we identified 12 exposures associated with gestational hypertension and 11 exposures associated with overall HDP. These exposures spanned diverse domains, including ambient air toxicants, blue space, neighborhood crime and safety, meteorological factors, environmental noise, and ultraviolet radiation. While the effect sizes were modest in magnitude and specific to subtype of HDP, our findings suggest that the total environment in which women live—in terms of natural, built, and social characteristics—represents a modifiable factor which could potentially influence their risk of gestational hypertension or HDP.

Our findings are broadly consistent with prior work linking spatial and contextual environmental exposures to HDP, while extending earlier exposome-wide studies through EHR-linked phenotyping, subtype-specific analyses, and multi-treatment estimation. In our prior ExWAS based on only vital statistics birth records in Florida,³² several air toxicants, meteorological variables, and neighborhood indicators were associated with HDP; we replicate key signals such as 2,2,4-trimethylpentane and neighborhood crime measures, supporting their robustness across data sources and analytic strategies. In addition to identifying

similar environmental signals in both of our approaches, the present study expands on our prior work to identify heterogeneity in effects by HDP subtype.

Among air toxicants, several compounds were positively associated with gestational hypertension and/or overall HDP in Phase 2. 2,2,4-trimethylpentane and catechol were consistently associated with higher odds of gestational hypertension, while 2,2,4-trimethylpentane, glycol ethers, and selenium compounds were associated with higher odds of overall HDP. These compounds likely reflect complex industrial or traffic-related mixtures and may influence HDP risk through oxidative stress, systemic inflammation, endothelial dysfunction, and placental vascular dysregulation—pathways strongly implicated in the pathophysiology of pregnancy hypertension.^{20,75,76} The strong association observed for selenium compounds is notable, as selenium is an essential micronutrient but may be harmful at elevated environmental levels; ambient selenium may also act as a marker of broader industrial emissions rather than a direct causal agent.⁷⁷⁻⁷⁹ Nevertheless, positive associations should be interpreted cautiously, as air toxicants often co-occur with other environmental and socioeconomic conditions, and residual confounding or correlated exposures may contribute to observed relationships despite adjustment. In contrast, several air toxicants showed inverse associations, including hydrogen fluoride (hydrofluoric acid) and methyl bromide (bromomethane) for both gestational hypertension and overall HDP, methyl isobutyl ketone (hexone) for gestational hypertension, and xylenes (mixed isomers) and 1,1,1-trichloroethane for overall HDP. These inverse associations should be interpreted cautiously. Similar paradoxical inverse associations have long been observed between maternal cigarette smoking and preeclampsia,^{80,81} where proposed biological mechanisms (eg, carbon monoxide-mediated effects on placental angiogenic balance) coexist with strong concerns about residual confounding and selection bias. In the exposome context, such inverse associations may reflect complex mixture effects, correlated unmeasured exposures, or differential measurement error rather than true protective effects.

Table 3. Odds ratios (ORs) and 95% confidence intervals (95% CIs) of gestational hypertension by exposome measures included in phase 2 analysis among singleton pregnancies with a conception date between 2013 and 2018 in the linked EHR-vital statistics birth and fetal death records data in Florida (n = 662,182).

Exposome measure	Category	Standard deviation	Phase 1 ^a discovery set			Phase 1 ^a replication set			Phase 2 ^b		
			OR (95% CI)	P-value	q-value	OR (95% CI)	P-value	q-value	OR (95% CI) ^c	P-value	
Hexane	Air toxicants	8.40 × 10 ⁻²	1.04 (1.02, 1.06)	1.76 × 10 ⁻⁶	5.01 × 10 ⁻⁵	1.04 (1.02, 1.05)	3.72 × 10 ⁻⁵	1.04 × 10 ⁻³	1.00 (0.98, 1.01)	5.52 × 10 ⁻¹	
Hydrogen fluoride (hydrofluoric acid)	Air toxicants	5.42 × 10 ⁻²	0.97 (0.96, 0.98)	8.34 × 10 ⁻⁸	3.19 × 10 ⁻⁶	0.98 (0.97, 0.99)	9.22 × 10 ⁻⁵	2.21 × 10 ⁻³	0.99 (0.98, 1.00)	2.94 × 10 ⁻²	
Methyl bromide (bromomethane)	Air toxicants	7.34 × 10 ⁻³	0.97 (0.96, 0.98)	1.58 × 10 ⁻⁶	4.61 × 10 ⁻⁵	0.97 (0.96, 0.99)	1.08 × 10 ⁻⁴	2.46 × 10 ⁻³	0.98 (0.98, 0.99)	2.47 × 10 ⁻⁴	
Methyl isobutyl ketone (hexone)	Air toxicants	3.75 × 10 ⁻²	1.03 (1.01, 1.05)	2.17 × 10 ⁻⁴	3.78 × 10 ⁻³	1.04 (1.02, 1.06)	1.86 × 10 ⁻⁶	6.15 × 10 ⁻⁵	0.97 (0.95, 1.00)	3.55 × 10 ⁻²	
2,2,4-trimethylpentane	Air toxicants	1.49 × 10 ⁻¹	1.06 (1.04, 1.08)	1.17 × 10 ⁻¹²	6.83 × 10 ⁻¹¹	1.03 (1.02, 1.05)	4.11 × 10 ⁻⁴	7.75 × 10 ⁻³	1.05 (1.03, 1.07)	8.66 × 10 ⁻⁷	
2,4-dinitrotoluene	Air toxicants	5.00 × 10 ⁻⁷	1.03 (1.01, 1.05)	1.36 × 10 ⁻³	1.89 × 10 ⁻²	1.03 (1.02, 1.05)	4.68 × 10 ⁻⁵	1.19 × 10 ⁻³	1.00 (0.99, 1.02)	7.10 × 10 ⁻¹	
Catechol	Air toxicants	5.08 × 10 ⁻⁴	1.05 (1.03, 1.06)	3.20 × 10 ⁻⁷	1.06 × 10 ⁻⁵	1.04 (1.02, 1.06)	9.79 × 10 ⁻⁴	1.62 × 10 ⁻²	1.03 (1.01, 1.05)	5.07 × 10 ⁻³	
Proximity to coastline (m)	Blue space	2.67 × 10 ⁴	1.06 (1.05, 1.07)	3.33 × 10 ⁻²⁴	3.40 × 10 ⁻²²	1.05 (1.04, 1.06)	1.98 × 10 ⁻¹⁵	1.17 × 10 ⁻¹³	1.04 (1.03, 1.05)	6.78 × 10 ⁻¹⁷	
Aggravated assault rate (per 100 population)	Crime and safety	9.62 × 10 ⁻²	1.03 (1.02, 1.04)	6.62 × 10 ⁻⁹	3.01 × 10 ⁻⁷	1.04 (1.03, 1.05)	1.37 × 10 ⁻¹⁰	6.73 × 10 ⁻⁹	1.00 (0.99, 1.02)	4.43 × 10 ⁻¹	
Burglary rate (per 100 population)	Crime and safety	2.00 × 10 ⁻¹	1.04 (1.03, 1.05)	4.65 × 10 ⁻¹¹	2.48 × 10 ⁻⁹	1.06 (1.05, 1.08)	1.46 × 10 ⁻¹⁹	1.19 × 10 ⁻¹⁷	1.02 (1.00, 1.04)	3.95 × 10 ⁻²	
Forcible sex offenses rate (per 100 population)	Crime and safety	1.53 × 10 ⁻²	1.05 (1.04, 1.06)	2.65 × 10 ⁻¹⁸	2.16 × 10 ⁻¹⁶	1.05 (1.03, 1.06)	5.10 × 10 ⁻¹³	2.84 × 10 ⁻¹¹	1.01 (1.00, 1.02)	1.43 × 10 ⁻²	
Larceny rate (per 100 population)	Crime and safety	6.43 × 10 ⁻¹	1.02 (1.01, 1.03)	4.36 × 10 ⁻⁴	6.93 × 10 ⁻³	1.04 (1.03, 1.05)	2.17 × 10 ⁻⁹	9.43 × 10 ⁻⁸	1.00 (0.96, 1.05)	8.22 × 10 ⁻¹	
Motor vehicle theft rate (per 100 population)	Crime and safety	8.70 × 10 ⁻²	1.02 (1.01, 1.03)	3.34 × 10 ⁻³	4.35 × 10 ⁻²	1.03 (1.01, 1.04)	1.23 × 10 ⁻⁴	2.74 × 10 ⁻³	0.99 (0.97, 1.01)	3.16 × 10 ⁻¹	
Robbery rate (per 100 population)	Crime and safety	5.55 × 10 ⁻²	1.02 (1.01, 1.03)	3.83 × 10 ⁻³	4.94 × 10 ⁻²	1.03 (1.01, 1.04)	9.66 × 10 ⁻⁵	2.28 × 10 ⁻³	1.01 (0.99, 1.03)	5.28 × 10 ⁻¹	
Total crime rate (per 100 population)	Crime and safety	9.78 × 10 ⁻¹	1.03 (1.02, 1.04)	1.22 × 10 ⁻⁷	4.40 × 10 ⁻⁶	1.04 (1.03, 1.06)	4.10 × 10 ⁻¹¹	2.18 × 10 ⁻⁹	0.99 (0.93, 1.05)	7.64 × 10 ⁻¹	
Percent vacant 36 months or longer	Land vacancy	1.75 × 10 ⁻²	0.99 (0.98, 1.00)	1.87 × 10 ⁻³	2.54 × 10 ⁻²	0.99 (0.98, 1.00)	2.82 × 10 ⁻³	3.80 × 10 ⁻²	1.00 (0.99, 1.01)	9.12 × 10 ⁻¹	
Heat index (deg C)	Meteorology	4.81	0.89 (0.87, 0.90)	9.11 × 10 ⁻⁴⁷	1.24 × 10 ⁻⁴⁴	0.86 (0.84, 0.87)	3.98 × 10 ⁻⁵³	6.09 × 10 ⁻⁵¹	0.97 (0.88, 1.07)	5.19 × 10 ⁻¹	
Maximum relative humidity (%)	Meteorology	3.39	1.05 (1.04, 1.06)	4.63 × 10 ⁻⁴⁵	5.67 × 10 ⁻⁴³	1.05 (1.04, 1.06)	7.58 × 10 ⁻³²	9.28 × 10 ⁻³⁰	1.02 (1.01, 1.04)	1.09 × 10 ⁻²	
Minimum relative humidity (%)	Meteorology	4.73	0.92 (0.92, 0.93)	4.17 × 10 ⁻⁷⁹	2.55 × 10 ⁻⁷⁶	0.92 (0.92, 0.93)	4.05 × 10 ⁻⁶¹	9.91 × 10 ⁻⁵⁹	0.95 (0.93, 0.98)	2.51 × 10 ⁻⁴	
Mean dew point temperature (deg C)	Meteorology	3.75	0.89 (0.88, 0.91)	2.14 × 10 ⁻⁴⁸	3.28 × 10 ⁻⁴⁶	0.86 (0.85, 0.88)	4.73 × 10 ⁻⁵⁸	8.28 × 10 ⁻⁵⁶	1.07 (0.96, 1.19)	2.07 × 10 ⁻¹	
Maximum temperature (deg C)	Meteorology	3.26	0.94 (0.93, 0.96)	7.44 × 10 ⁻¹³	4.56 × 10 ⁻¹¹	0.92 (0.90, 0.94)	3.26 × 10 ⁻¹⁹	2.50 × 10 ⁻¹⁷	0.92 (0.86, 1.00)	3.73 × 10 ⁻²	
Minimum temperature (deg C)	Meteorology	4.40	0.90 (0.89, 0.91)	4.68 × 10 ⁻⁶²	1.43 × 10 ⁻⁵⁹	0.89 (0.88, 0.90)	5.90 × 10 ⁻⁶⁰	1.20 × 10 ⁻⁵⁷	1.04 (0.95, 1.12)	4.05 × 10 ⁻¹	
Impact A-weighted L50 sound pressure level, 1230 m buffer	Noise	3.09	1.04 (1.03, 1.06)	5.90 × 10 ⁻⁸	2.33 × 10 ⁻⁶	1.03 (1.01, 1.04)	3.15 × 10 ⁻⁴	6.03 × 10 ⁻³	0.98 (0.95, 1.01)	2.62 × 10 ⁻¹	
Impact A-weighted L50 sound pressure level, 270 m buffer	Noise	3.00	1.03 (1.02, 1.05)	6.89 × 10 ⁻⁷	2.11 × 10 ⁻⁵	1.02 (1.01, 1.03)	3.12 × 10 ⁻³	4.13 × 10 ⁻²	1.01 (0.98, 1.04)	5.09 × 10 ⁻¹	
Natural A-weighted L50 sound pressure level, 1230 m buffer	Noise	8.15 × 10 ⁻¹	0.94 (0.93, 0.96)	1.96 × 10 ⁻²⁰	1.71 × 10 ⁻¹⁸	0.94 (0.93, 0.95)	5.16 × 10 ⁻¹⁹	3.51 × 10 ⁻¹⁷	0.99 (0.98, 1.01)	2.96 × 10 ⁻¹	
Erythral UV index	Ultra violet	2.09	0.86 (0.79, 0.94)	1.14 × 10 ⁻³	1.65 × 10 ⁻²	0.83 (0.75, 0.92)	5.93 × 10 ⁻⁴	1.07 × 10 ⁻²	0.94 (0.90, 0.98)	4.39 × 10 ⁻³	

^aPhase 1 models adjusted for different confounders specific to each exposome measure as shown in Table S8

^bPhase 2 adjusted for race/ethnicity, education, marital status, insurance type, parity, month of conception, year of conception, road proximity, green space, walkability, and neighborhood deprivation.

^cStatistically significant results in bold.

Table 4. Odds ratios (ORs) and 95% confidence intervals (95% CIs) of overall HDP by exposome measures included in phase 2 analysis among singleton pregnancies with a conception date between 2013 and 2018 in the linked EHR-vital statistics birth and fetal death records data in Florida (n = 686,412).

Exposure	Phase 1 ^a discovery set				Phase 1 ^a replication set				Phase 2 ^b			
	Standard deviation	Or (95% CI)	P-value	q-value	Or (95% CI)	P-value	q-value	Or (95% CI) ^c	P-value	q-value	Or (95% CI) ^c	P-value
Ethylene glycol	1.88 × 10 ⁻¹	1.04 (1.02, 1.06)	1.65 × 10 ⁻⁵	3.55 × 10 ⁻⁴	1.04 (1.02, 1.06)	6.30 × 10 ⁻⁵	1.58 × 10 ⁻³	0.98 (0.95, 1.00)	6.66 × 10 ⁻²	0.98 (0.95, 1.00)	6.66 × 10 ⁻²	0.98 (0.95, 1.00)
Glycol ethers	4.26 × 10 ⁻²	1.04 (1.02, 1.06)	6.24 × 10 ⁻⁶	1.93 × 10 ⁻⁴	1.03 (1.01, 1.05)	4.53 × 10 ⁻⁴	8.27 × 10 ⁻³	1.03 (1.00, 1.05)	3.84 × 10 ⁻²	1.03 (1.00, 1.05)	3.84 × 10 ⁻²	1.03 (1.00, 1.05)
Air toxicants	8.40 × 10 ⁻²	1.04 (1.03, 1.06)	2.26 × 10 ⁻⁷	7.93 × 10 ⁻⁶	1.04 (1.02, 1.05)	3.04 × 10 ⁻⁵	8.87 × 10 ⁻⁴	1.02 (1.00, 1.04)	5.09 × 10 ⁻²	1.02 (1.00, 1.04)	5.09 × 10 ⁻²	1.02 (1.00, 1.04)
Hexane	5.42 × 10 ⁻²	0.98 (0.97, 0.99)	1.29 × 10 ⁻⁴	2.32 × 10 ⁻³	0.98 (0.97, 0.99)	2.25 × 10 ⁻⁴	4.45 × 10 ⁻³	0.99 (0.98, 0.99)	5.38 × 10 ⁻⁴	0.99 (0.98, 0.99)	5.38 × 10 ⁻⁴	0.99 (0.98, 0.99)
Hydrogen fluoride (hydrofluoric acid)	7.34 × 10 ⁻³	0.98 (0.97, 0.99)	6.28 × 10 ⁻⁵	1.18 × 10 ⁻³	0.97 (0.96, 0.98)	8.57 × 10 ⁻⁶	2.62 × 10 ⁻⁴	0.99 (0.98, 0.99)	6.17 × 10 ⁻⁴	0.99 (0.98, 0.99)	6.17 × 10 ⁻⁴	0.99 (0.98, 0.99)
Methyl bromide (bromomethane)	3.75 × 10 ⁻²	1.04 (1.03, 1.06)	1.18 × 10 ⁻⁶	3.52 × 10 ⁻⁵	1.06 (1.04, 1.07)	2.23 × 10 ⁻⁹	9.43 × 10 ⁻⁸	1.01 (0.98, 1.03)	6.31 × 10 ⁻¹	1.01 (0.98, 1.03)	6.31 × 10 ⁻¹	1.01 (0.98, 1.03)
Methyl isobutyl ketone (hexone)	4.65 × 10 ⁻⁷	0.98 (0.97, 0.99)	7.56 × 10 ⁻⁴	1.14 × 10 ⁻²	0.98 (0.98, 0.99)	2.37 × 10 ⁻³	3.26 × 10 ⁻²	1.00 (0.99, 1.00)	3.32 × 10 ⁻¹	1.00 (0.99, 1.00)	3.32 × 10 ⁻¹	1.00 (0.99, 1.00)
Quinone (p-benzoquinone)	5.62 × 10 ⁻⁵	1.02 (1.01, 1.03)	1.15 × 10 ⁻³	1.65 × 10 ⁻²	1.03 (1.02, 1.04)	3.37 × 10 ⁻⁵	9.60 × 10 ⁻⁴	1.04 (1.03, 1.05)	3.65 × 10 ⁻²¹	1.04 (1.03, 1.05)	3.65 × 10 ⁻²¹	1.04 (1.03, 1.05)
Selenium compounds	4.32 × 10 ⁻¹	1.04 (1.02, 1.06)	4.52 × 10 ⁻⁶	1.23 × 10 ⁻⁴	1.04 (1.02, 1.06)	4.15 × 10 ⁻⁵	1.08 × 10 ⁻³	1.02 (0.98, 1.05)	3.02 × 10 ⁻¹	1.02 (0.98, 1.05)	3.02 × 10 ⁻¹	1.02 (0.98, 1.05)
Toluene	2.85 × 10 ⁻¹	1.04 (1.02, 1.06)	2.63 × 10 ⁻⁴	4.48 × 10 ⁻³	1.04 (1.02, 1.06)	1.58 × 10 ⁻⁴	3.34 × 10 ⁻³	0.95 (0.92, 0.98)	1.57 × 10 ⁻³	0.95 (0.92, 0.98)	1.57 × 10 ⁻³	0.95 (0.92, 0.98)
Xylenes (mixed isomers)	1.49 × 10 ⁻¹	1.06 (1.04, 1.07)	7.24 × 10 ⁻¹¹	4.56 × 10 ⁻¹¹	1.04 (1.03, 1.06)	6.24 × 10 ⁻⁷	2.18 × 10 ⁻⁵	1.04 (1.02, 1.06)	3.18 × 10 ⁻⁶	1.04 (1.02, 1.06)	3.18 × 10 ⁻⁶	1.04 (1.02, 1.06)
2,2,4-trimethylpentane	1.92 × 10 ⁻²	1.05 (1.03, 1.06)	1.96 × 10 ⁻⁸	8.27 × 10 ⁻⁷	1.04 (1.02, 1.07)	1.28 × 10 ⁻⁴	2.81 × 10 ⁻³	0.97 (0.95, 0.99)	1.38 × 10 ⁻²	0.97 (0.95, 0.99)	1.38 × 10 ⁻²	0.97 (0.95, 0.99)
1,1,1-trichloroethane	5.00 × 10 ⁻⁷	1.04 (1.02, 1.06)	4.61 × 10 ⁻⁵	8.96 × 10 ⁻⁴	1.03 (1.02, 1.04)	3.95 × 10 ⁻⁵	1.05 × 10 ⁻³	1.01 (1.00, 1.02)	1.81 × 10 ⁻¹	1.01 (1.00, 1.02)	1.81 × 10 ⁻¹	1.01 (1.00, 1.02)
2,4-dinitrotoluene	5.96 × 10 ⁻⁵	1.04 (1.02, 1.06)	6.01 × 10 ⁻⁶	1.53 × 10 ⁻⁴	1.04 (1.02, 1.06)	3.86 × 10 ⁻⁵	1.05 × 10 ⁻³	0.99 (0.97, 1.01)	4.88 × 10 ⁻¹	0.99 (0.97, 1.01)	4.88 × 10 ⁻¹	0.99 (0.97, 1.01)
4-nitrophenol	5.08 × 10 ⁻⁴	1.06 (1.04, 1.08)	1.02 × 10 ⁻¹¹	5.67 × 10 ⁻¹⁰	1.06 (1.03, 1.08)	9.45 × 10 ⁻⁷	3.21 × 10 ⁻⁵	1.02 (1.00, 1.04)	5.82 × 10 ⁻²	1.02 (1.00, 1.04)	5.82 × 10 ⁻²	1.02 (1.00, 1.04)
Catechol	2.67 × 10 ⁴	1.05 (1.04, 1.06)	1.45 × 10 ⁻²²	1.37 × 10 ⁻²⁰	1.05 (1.04, 1.06)	4.51 × 10 ⁻¹⁹	3.25 × 10 ⁻¹⁷	1.02 (1.02, 1.03)	6.66 × 10 ⁻⁸	1.02 (1.02, 1.03)	6.66 × 10 ⁻⁸	1.02 (1.02, 1.03)
Proximity to coastline (m)	9.62 × 10 ⁻²	1.03 (1.02, 1.04)	1.01 × 10 ⁻⁹	4.74 × 10 ⁻⁸	1.05 (1.04, 1.07)	8.02 × 10 ⁻¹⁹	5.17 × 10 ⁻¹⁷	1.01 (1.00, 1.02)	9.74 × 10 ⁻²	1.01 (1.00, 1.02)	9.74 × 10 ⁻²	1.01 (1.00, 1.02)
Aggravated assault rate (per 100 population)	2.00 × 10 ⁻¹	1.04 (1.03, 1.05)	5.33 × 10 ⁻¹³	3.63 × 10 ⁻¹¹	1.07 (1.06, 1.08)	5.79 × 10 ⁻²⁷	5.91 × 10 ⁻²⁵	1.01 (1.00, 1.03)	5.84 × 10 ⁻²	1.01 (1.00, 1.03)	5.84 × 10 ⁻²	1.01 (1.00, 1.03)
Burglary rate (per 100 population)	1.53 × 10 ⁻²	1.05 (1.03, 1.06)	8.22 × 10 ⁻¹⁷	5.92 × 10 ⁻¹⁵	1.06 (1.04, 1.07)	4.45 × 10 ⁻²³	3.89 × 10 ⁻²¹	1.01 (1.00, 1.02)	6.36 × 10 ⁻²	1.01 (1.00, 1.02)	6.36 × 10 ⁻²	1.01 (1.00, 1.02)
Forcible sex offenses rate (per 100 population)	6.43 × 10 ⁻¹	1.02 (1.01, 1.03)	5.04 × 10 ⁻⁴	7.92 × 10 ⁻³	1.04 (1.02, 1.05)	1.18 × 10 ⁻⁹	5.38 × 10 ⁻⁸	0.99 (0.97, 1.01)	2.65 × 10 ⁻¹	0.99 (0.97, 1.01)	2.65 × 10 ⁻¹	0.99 (0.97, 1.01)
Larceny rate (per 100 population)	2.61 × 10 ⁻³	1.02 (1.01, 1.03)	7.87 × 10 ⁻⁴	1.18 × 10 ⁻²	1.04 (1.02, 1.05)	1.51 × 10 ⁻¹⁰	7.11 × 10 ⁻⁹	1.01 (1.00, 1.01)	1.41 × 10 ⁻¹	1.01 (1.00, 1.01)	1.41 × 10 ⁻¹	1.01 (1.00, 1.01)
Murder rate (per 100 population)	9.78 × 10 ⁻¹	1.02 (1.01, 1.04)	6.19 × 10 ⁻⁶	1.53 × 10 ⁻⁴	1.05 (1.04, 1.06)	2.00 × 10 ⁻¹⁵	1.17 × 10 ⁻¹³	1.00 (0.98, 1.03)	7.63 × 10 ⁻¹	1.00 (0.98, 1.03)	7.63 × 10 ⁻¹	1.00 (0.98, 1.03)
Total crime rate (per 100 population)	6.20 × 10 ²	0.98 (0.98, 0.99)	1.31 × 10 ⁻⁵	2.91 × 10 ⁻⁴	0.99 (0.98, 1.00)	1.58 × 10 ⁻³	2.34 × 10 ⁻²	1.00 (0.99, 1.00)	4.66 × 10 ⁻¹	1.00 (0.99, 1.00)	4.66 × 10 ⁻¹	1.00 (0.99, 1.00)
Average days addresses vacant	1.75 × 10 ⁻²	0.98 (0.98, 0.99)	1.26 × 10 ⁻⁵	2.87 × 10 ⁻⁴	0.98 (0.97, 0.99)	3.62 × 10 ⁻⁷	1.39 × 10 ⁻⁵	1.00 (0.99, 1.00)	3.01 × 10 ⁻¹	1.00 (0.99, 1.00)	3.01 × 10 ⁻¹	1.00 (0.99, 1.00)
Percent vacant 36 months or longer	4.81	0.88 (0.86, 0.89)	1.34 × 10 ⁻⁵²	2.74 × 10 ⁻⁵⁰	0.84 (0.83, 0.86)	1.34 × 10 ⁻⁶⁵	4.12 × 10 ⁻⁶³	1.02 (0.96, 1.09)	5.15 × 10 ⁻¹	1.02 (0.96, 1.09)	5.15 × 10 ⁻¹	1.02 (0.96, 1.09)
Heat index (deg C)	3.39	1.05 (1.05, 1.06)	1.39 × 10 ⁻⁴⁸	2.44 × 10 ⁻⁴⁶	1.05 (1.04, 1.06)	9.53 × 10 ⁻³⁷	1.30 × 10 ⁻³⁴	1.02 (1.00, 1.03)	3.39 × 10 ⁻²	1.02 (1.00, 1.03)	3.39 × 10 ⁻²	1.02 (1.00, 1.03)
Maximum relative humidity (%)	4.73	0.92 (0.91, 0.93)	3.25 × 10 ⁻⁸¹	3.98 × 10 ⁻⁷⁸	0.92 (0.91, 0.93)	6.97 × 10 ⁻⁷²	4.27 × 10 ⁻⁶⁹	1.00 (0.98, 1.02)	7.58 × 10 ⁻¹	1.00 (0.98, 1.02)	7.58 × 10 ⁻¹	1.00 (0.98, 1.02)
Minimum relative humidity (%)	3.75	0.88 (0.87, 0.90)	9.20 × 10 ⁻⁵⁵	2.25 × 10 ⁻⁵²	0.85 (0.83, 0.87)	3.04 × 10 ⁻⁶⁹	1.24 × 10 ⁻⁶⁶	0.94 (0.88, 1.00)	6.93 × 10 ⁻²	0.94 (0.88, 1.00)	6.93 × 10 ⁻²	0.94 (0.88, 1.00)
Mean dew point temperature (deg C)	3.26	0.93 (0.92, 0.95)	5.22 × 10 ⁻¹⁸	3.99 × 10 ⁻¹⁶	0.91 (0.89, 0.93)	2.68 × 10 ⁻²⁴	2.52 × 10 ⁻²²	1.03 (0.97, 1.08)	3.31 × 10 ⁻¹	1.03 (0.97, 1.08)	3.31 × 10 ⁻¹	1.03 (0.97, 1.08)
Maximum temperature (deg C)	4.40	0.89 (0.88, 0.90)	9.89 × 10 ⁻⁶⁷	4.04 × 10 ⁻⁶⁴	0.88 (0.87, 0.89)	1.52 × 10 ⁻⁷²	1.87 × 10 ⁻⁶⁹	1.01 (0.96, 1.07)	6.78 × 10 ⁻¹	1.01 (0.96, 1.07)	6.78 × 10 ⁻¹	1.01 (0.96, 1.07)
Minimum temperature (deg C)	3.09	1.05 (1.03, 1.06)	1.33 × 10 ⁻¹⁰	6.53 × 10 ⁻⁹	1.04 (1.02, 1.05)	7.04 × 10 ⁻⁸	2.78 × 10 ⁻⁶	1.00 (0.97, 1.02)	8.47 × 10 ⁻¹	1.00 (0.97, 1.02)	8.47 × 10 ⁻¹	1.00 (0.97, 1.02)
Impact A-weighted L50 sound pressure level, 1230m buffer	3.00	1.04 (1.03, 1.06)	5.82 × 10 ⁻¹¹	2.97 × 10 ⁻⁹	1.03 (1.02, 1.05)	2.28 × 10 ⁻⁶	7.16 × 10 ⁻⁵	0.99 (0.97, 1.02)	6.42 × 10 ⁻¹	0.99 (0.97, 1.02)	6.42 × 10 ⁻¹	0.99 (0.97, 1.02)
Impact A-weighted L50 sound pressure level, 270m buffer	8.15 × 10 ⁻¹	0.94 (0.92, 0.95)	2.14 × 10 ⁻²⁵	2.39 × 10 ⁻²³	0.93 (0.92, 0.94)	2.86 × 10 ⁻³¹	3.19 × 10 ⁻²⁹	0.98 (0.97, 0.99)	1.19 × 10 ⁻³	0.98 (0.97, 0.99)	1.19 × 10 ⁻³	0.98 (0.97, 0.99)
Natural A-weighted L50 sound pressure level, 1230m buffer	2.09	0.86 (0.79, 0.94)	7.32 × 10 ⁻⁴	1.12 × 10 ⁻²	0.84 (0.76, 0.92)	2.23 × 10 ⁻⁴	4.45 × 10 ⁻³	0.94 (0.91, 0.98)	2.11 × 10 ⁻³	0.94 (0.91, 0.98)	2.11 × 10 ⁻³	0.94 (0.91, 0.98)
Erythral UV index												

^aPhase 1 models adjusted for different confounders specific to each exposome measure as shown in Table S12

^bPhase 2 adjusted for race/ethnicity, education, marital status, insurance type, parity, month of conception, year of conception, road proximity, green space, walkability, and neighborhood deprivation.

^cStatistically significant results in bold.

Differences from prior studies may also reflect variation in spatial scale, exposure definitions, regional emission profiles, or analytic approaches. Accordingly, these findings should be viewed as hypothesis-generating.

Meteorological and ultraviolet exposures also showed consistent and biologically plausible patterns. Maximum relative humidity was positively associated with both gestational hypertension and overall HDP, whereas minimum relative humidity and maximum temperature were inversely associated with gestational hypertension. In addition, higher erythemal UV index was associated with lower odds of both gestational hypertension and overall HDP. These findings are consistent with prior evidence of seasonal variation in HDP and hypotheses linking sunlight exposure, vitamin D metabolism, and vascular function.⁸²⁻⁸⁷

Several built and social environment measures were independently associated with HDP outcomes. Living closer to the coastline was associated with lower odds of both gestational hypertension and overall HDP, potentially reflecting broader coastal or blue space-related environmental and contextual characteristics rather than a single etiologic factor.⁸⁸ In the crime and safety domain, higher burglary rates and forcible sex offenses rates were associated with increased odds of gestational hypertension, aligning with prior literature linking neighborhood stressors to elevated risk of preeclampsia and gestational hypertension.^{25,32,89,90} Chronic psychosocial stress related to neighborhood safety may contribute to dysregulation of the hypothalamic-pituitary-adrenal axis and vascular function,⁹¹ thereby increasing susceptibility to gestational hypertension.^{92,93} Similarly, environmental noise exposure—another stress-related urban hazard—has been implicated in HDP. While impact-weighted noise measures did not remain significant after mutual adjustment, higher natural A-weighted L50 sound pressure levels were inversely associated with overall HDP, suggesting that quieter, less anthropogenically disturbed environments may be beneficial for maternal cardiovascular health. This finding is consistent with growing evidence linking chronic noise exposure to hypertension and cardiometabolic risk.^{94,95}

A central motivation for our study was to examine whether different HDP subtypes have distinct environmental risk factor profiles. Indeed, our findings suggest notable heterogeneity between gestational hypertension and the preeclampsia spectrum. Gestational hypertension—often considered a more benign or milder condition—showed the richest exposome signal, with 12 exposures replicating and remaining independently associated in Phase 2. In contrast, preeclampsia and eclampsia had no exposure passing our discovery and replication thresholds. These differences could have several explanations. First and mostly likely, due to different sample sizes for HDP subtypes, the statistical power to detect exposome-wide associations was much greater for gestational hypertension than for severe preeclampsia or eclampsia. Therefore, it is possible that meaningful but modest associations for severe preeclampsia were missed. Second, there may be true etiological differences: gestational hypertension may be more influenced by environmental stressors, whereas severe preeclampsia could be more strongly driven by placental genetic/immune factors that are less tied to the external environment. This concept has been supported by the literature: eg, one study found air pollution exposure was associated with higher risk of mild preeclampsia but not with severe disease.⁹⁶ It is also possible that our stringent analysis (controlling false discovery and using replication) was too conservative for the smaller subtype strata. For example, if we relaxed the multiple-testing correction,

we might find that severe preeclampsia was nominally associated with three exposome measures (ie, minimum relative humidity, natural A-weighted L50 sound pressure level and arsenic compounds). Our inability to detect such links in an agnostic screen likely reflects a combination of limited power and the challenge of multiple comparisons. Our results highlight that future studies should stratify HDP by subtype to unmask associations that might be diluted in composite outcomes. Larger exposome studies or targeted environmental analyses for severe preeclampsia may eventually identify exposures that we were underpowered to detect.

This study has several strengths. First, we combined statewide vital statistics with EHR data to improve both scale and phenotype specificity, enabling ascertainment of HDP subtypes and reconstruction of pregnancy residential histories. We evaluated a broad set of spatial and contextual exposome measures across multiple domains and used a rigorous two-phase design (temporal discovery/replication followed by multi-treatment estimation) to reduce false positives and identify exposome measures independently associated with the outcomes after accounting for co-exposures. Finally, our analytic framework was grounded in causal inference principles and modern machine learning. We departed from the standard exposome study practice of using one-size-fits-all covariate adjustment for every exposure.³⁰ Instead, we developed exposure-specific confounder sets guided by DAG, recognizing that different exposures have different confounding structures. To handle the high dimensionality of covariates and account for potential non-linearities, we applied DML for effect estimation, which uses machine learning to flexibly model nuisance variables, and then estimates the effect size orthogonal to those nuisances. This approach provides efficient adjustment for many confounders without manual model specification, and importantly, yields asymptotically unbiased effect estimates with valid confidence intervals under moderate assumptions. To our knowledge, this is the first spatial and contextual exposome study to apply a formal DML approach. The successful implementation of DML in our study demonstrates its feasibility for exposome analyses, addressing a key analytic challenge in this field: how to adjust for the dense web of correlated covariates and exposures without overfitting or introducing bias. We see this as a methodological contribution that can be extended to other exposome-health investigations to improve causal inference rigor.

Several limitations should be considered. First, exposure assignment relied on secondary, model-based or contextual-level measures, which may not fully capture individual time-activity patterns and can introduce measurement error (likely nondifferential, and therefore potentially biasing effects toward the null). Second, residential histories were derived from encounter-based EHR address records; while address start and end dates were available, they may not reliably reflect true move-in and move-out dates, and address update frequency may vary across individuals. We did not have information on workplace locations or commuting patterns; thus, exposures experienced outside the home environment were not captured. In addition, address histories were derived from in-state EHR records and may not capture periods of residence outside Florida. Third, housing instability is not captured in our data, which may introduce exposure misclassification for a small subset of women. Fourth, although EHR linkage improved outcome definitions, some misclassification likely remains (eg, undiagnosed chronic hypertension or incomplete documentation). While sensitivity analyses further adjusting for residential mobility and urbanicity yielded similar effect

estimates, residual confounding cannot be excluded—particularly for individual behaviors and health status factors that may be imperfectly captured. Fifth, generalizability may be limited to Florida’s climate, geography, and healthcare context, and selection into the linked EHR network could introduce bias. Finally, stringent multiple-comparison control and replication improve credibility but may increase false negatives, especially for rarer subtypes (eg, eclampsia).

Conclusions

In conclusion, our exposome-wide analysis of a statewide pregnancy cohort identified a set of diverse environmental factors—ranging from air toxicants and meteorology to neighborhood built and social characteristics—that are independently associated with HDP. These findings support the premise that the spatial and contextual exposome contributes to HDP risk and highlight modifiable domains that could inform prevention strategies.

Author contributions

Hui Hu (Conceptualization [lead], Funding acquisition [lead], Formal analysis [lead], Methodology [lead], Visualization [lead], Writing-original draft [lead], Writing-review & editing [lead]), Claire Leiser (Writing-review & editing [supporting]), Xing He (Writing-review & editing [supporting]), Jaime Hart (Writing-review & editing [supporting]), Francine Laden (Writing-review & editing [supporting]), Cui Tao (Writing-review & editing [supporting]), and Jiang Bian (Writing-review & editing [supporting])

Funding

Research reported in this publication was supported in part by the National Heart, Lung, and Blood Institute under award number K01HL153797; in part by the National Institute of Environmental Health Sciences under award number R24ES036131 and P30ES000002; in part by OneFlorida+ Clinical Research Network, funded by the Patient-Centered Outcomes Research Institute numbers CDRN-1501-26692 and RI-CRN-2020-005; in part by the OneFlorida Cancer Control Alliance, funded by the Florida Department of Health’s James and Esther King Biomedical Research Program number 4KB16; and in part by the University of Florida Clinical and Translational Science Institute and its Clinical and Translational Science Award (CTSA) hub partner, Florida State University (FSU), which are supported in part by the National Center for Advancing Translational Sciences of the National Institutes of Health grant numbers UL1TR001427, KL2TR001429, and TL1TR001428. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology, the OneFlorida+ Clinical Research Network, the UF-FSU Clinical and Translational Science Institute, the Florida Department of Health, or the National Institutes of Health.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

The data underlying this article were provided by the OneFlorida+ Clinical Research Network (<https://onefloridaconsortium.org/>), which are made available to researchers with an approved study protocol and data use agreement at <https://onefloridaconsortium.org/front-door/prep-to-research-data-query/>. The data will be shared on reasonable requests to the corresponding author with the permission of OneFlorida+ Clinical Research Network.

References

1. Radparvar AA, Vani K, Fiori K, et al. Hypertensive disorders of pregnancy: innovative management strategies. *JACC Adv.* 2024; 3(3):100864.
2. Duley L. The global impact of pre-eclampsia and eclampsia. *Semin Perinatol.* 2009;33(3):130–137.
3. Lo JO, Mission JF, Caughey AB. Hypertensive disease of pregnancy and maternal mortality. *Curr Opin Obstet Gynecol.* 2013;25(2):124–132.
4. Bauer ST, Cleary KL. Cardiopulmonary complications of pre-eclampsia. *Semin Perinatol.* 2009;33(3):158–165.
5. Bellamy L, Casas JP, Hingorani AD, Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *Bmj.* 2007;335(7627):974.
6. Wang IK, Tsai IJ, Chen PC, et al. Hypertensive disorders in pregnancy and subsequent diabetes mellitus: a retrospective cohort study. *Am J Med.* 2012;125(3):251–257.
7. Allen VM, Joseph KS, Murphy KE, Magee LA, Ohlsson A. The effect of hypertensive disorders in pregnancy on small for gestational age and stillbirth: a population based study. *BMC Pregnancy Childbirth.* 2004;4(1):17.
8. Wu CS, Nohr EA, Bech BH, Vestergaard M, Catov JM, Olsen J. Health of children born to mothers who had preeclampsia: a population-based cohort study. *Am J Obstet Gynecol.* 2009;201(3): 269.e1–269.e10.
9. Premkumar A, Baer RJ, Jelliffe-Pawlowski LL, Norton ME. Hypertensive disorders of pregnancy and preterm birth rates among black women. *Am J Perinatol.* 2019;36(2):148–154.
10. Di Martino DD, Avagliano L, Ferrazzi E, et al. Hypertensive disorders of pregnancy and fetal growth restriction: clinical characteristics and placental lesions and possible preventive nutritional targets. *Nutrients* 2022;14(16):3276.
11. Li F, Wang T, Chen L, Zhang S, Chen L, Qin J. Adverse pregnancy outcomes among mothers with hypertensive disorders in pregnancy: a meta-analysis of cohort studies. *Pregnancy Hypertens.* 2021;24:107–117.
12. Wu P, Green M, Myers JE. Hypertensive disorders of pregnancy. *BMJ.* 2023;381:e071653.
13. Garovic VD, Bailey KR, Boerwinkle E, et al. Hypertension in pregnancy as a risk factor for cardiovascular disease later in life. *J Hypertens.* 2010;28(4):826–833.
14. Lykke JA, Langhoff-Roos J, Sibai BM, Funai EF, Triche EW, Paidas MJ. Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother. *Hypertension.* 2009;53(6):944–951.
15. Tooher J, Thornton C, Makris A, et al. Hypertension in pregnancy and long-term cardiovascular mortality: a retrospective cohort study. *Am J Obstet Gynecol.* 2016;214(6):722.e1–6–722.e6.
16. Brown MC, Best KE, Pearce MS, Waugh J, Robson SC, Bell R. Cardiovascular disease risk in women with pre-eclampsia:

- systematic review and meta-analysis. *Eur J Epidemiol*. 2013; 28(1):1–19.
17. McDonald SD, Malinowski A, Zhou Q, Yusuf S, Devereaux PJ. Cardiovascular sequelae of preeclampsia/eclampsia: a systematic review and meta-analyses. *Am Heart J*. 2008;156(5):918–930.
 18. Hu H, Xiao H, Zheng Y, Yu BB. A Bayesian spatio-temporal analysis on racial disparities in hypertensive disorders of pregnancy in Florida, 2005–2014. *Spat Spatiotemporal Epidemiol*. 2019; 29:43–50.
 19. World Health Organization International Collaborative Study of Hypertensive Disorders of Pregnancy. Geographic variation in the incidence of hypertension in pregnancy. *Am J Obstet Gynecol*. 1988;158(1):80–83.
 20. Hu H, Ha S, Roth J, Kearney G, Talbott EO, Xu X. Ambient air pollution and hypertensive disorders of pregnancy: a systematic review and meta-analysis. *Atmos Environ*. 2014;97:336–345.
 21. Osorio-Yañez C, Gelaye B, Miller RS, et al. Associations of maternal urinary cadmium with trimester-specific blood pressure in pregnancy: Role of dietary intake of micronutrients. *Biol Trace Elem Res*. 2016;174(1):71–81.
 22. Strand LB, Barnett AG, Tong S. The influence of season and ambient temperature on birth outcomes: a review of the epidemiological literature. *Environ Res*. 2011;111(3):451–462.
 23. Tran TC, Boumendil A, Bussieres L, et al. Are meteorological conditions within the first trimester of pregnancy associated with the risk of severe pre-eclampsia? *Paediatr Perinat Epidemiol*. 2015;29(4):261–270.
 24. Vinikoor-Imler LC, Gray SC, Edwards SE, Miranda ML. The effects of exposure to particulate matter and neighbourhood deprivation on gestational hypertension. *Paediatr Perinat Epidemiol*. 2012;26(2):91–100.
 25. Mayne SL, Pool LR, Grobman WA, Kershaw KN. Associations of neighbourhood crime with adverse pregnancy outcomes among women in Chicago: analysis of electronic health records from 2009 to 2013. *J Epidemiol Community Health*. 2018;72(3):230–236.
 26. Messer LC, Vinikoor-Imler LC, Laraia BA. Conceptualizing neighborhood space: consistency and variation of associations for neighborhood factors and pregnancy health across multiple neighborhood units. *Health Place*. 2012;18(4):805–813.
 27. Grazuleviciene R, Dedele A, Danileviciute A, et al. The influence of proximity to city parks on blood pressure in early pregnancy. *Int J Environ Res Public Health*. 2014;11(3):2958–2972.
 28. Morales ME, Epstein MH, Marable DE, Oo SA, Berkowitz SA. Food insecurity and cardiovascular health in pregnant women: results from the food for Families program, Chelsea, Massachusetts, 3–2015. *Prev Chronic Dis*. 2016;13:E152.
 29. Wild CP. The exposome: from concept to utility. *Int J Epidemiol*. 2012;41(1):24–32.
 30. Hu H, Liu X, Zheng Y, et al. Methodological challenges in spatial and contextual exposome-health studies. *Crit Rev Environ Sci Technol*. 2023;53(7):827–846. Published online July 4
 31. Hu H, Laden F, Hart J, et al. A spatial and contextual exposome-wide association study and polyexposomic score of COVID-19 hospitalization. *Exposome* 2023;3(1):osad005. <https://doi.org/10.1093/exposome/osad005>
 32. Hu H, Zhao J, Savitz DA, Prosperi M, Zheng Y, Pearson TA. An external exposome-wide association study of hypertensive disorders of pregnancy. *Environ Int*. 2020;141:105797.
 33. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011;25(4):391–403.
 34. Trogstad L, Magnus P, Stoltenberg C. Pre-eclampsia: risk factors and causal models. *Best Pract Res Clin Obstet Gynaecol*. 2011; 25(3):329–342.
 35. Dadvand P, Figueras F, Basagaña X, et al. Ambient air pollution and preeclampsia: a spatiotemporal analysis. *Environ Health Perspect*. 2013;121(1–12):1365–1371.
 36. Mendola P, Wallace M, Liu D, Robledo C, Männistö T, Grantz KL. Air pollution exposure and preeclampsia among US women with and without asthma. *Environ Res*. 2016; 148:248–255.
 37. Bell ML, Belanger K. Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. *J Expo Sci Environ Epidemiol*. 2012;22 (5):429–438.
 38. Hodgson S, Lurz PWW, Shirley MDF, Bythell M, Rankin J. Exposure misclassification due to residential mobility during pregnancy. *Int J Hyg Environ Health*. 2015;218(4):414–421.
 39. Pennington AF, Strickland MJ, Klein M, et al. Measurement error in mobile source air pollution exposure estimates due to residential mobility during pregnancy. *J Expo Sci Environ Epidemiol*. 2017;27(5):513–520.
 40. Al-Taai SHH, Mohammed Al-Dulaimi WA. Air pollution: a study of its concept, causes, sources and effects. *Ajwep*. 2022;19 (1):17–22.
 41. Bilgili BC, Gökyer E. Urban green space system planning. *Landsc Plan*. 2012;360:107–123.
 42. Lipsky AM, Greenland S. Causal directed acyclic graphs. *JAMA* 2022;327(11):1083–1084.
 43. Shiba K, Kawahara T. Using propensity scores for causal inference: Pitfalls and tips. *J Epidemiol*. 2021;31(8):457–463.
 44. Miguel A, Hernan R, James M. Causal inference: what if. 2023.
 45. Hernán MA, Robins JM. Causal inference. 2010. https://grass.upc.edu/en/seminar/presentation-files/causal-inference/chapters-1-i-2/@@download/file/BookHernanRobinsCap1_2.pdf
 46. Chernozhukov V, Chetverikov D, Demirem M, et al. Double/debiased machine learning for treatment and structural parameters. *Econom J*. 2018;21(1):C1–C68.
 47. Chernozhukov V, Chetverikov D, Demirem M, Duflo E, Hansen C, Newey W. Double/debiased/Neyman machine learning of treatment effects. *Am Econ Rev*. 2017;107(5):261–265.
 48. Shenkman E, Hurt M, Hogan W, et al. OneFlorida Clinical Research Consortium: Linking a clinical and Translational Science Institute with a community-based distributive medical education model. *Acad Med*. 2018;93(3):451–455.
 49. Narang K, Szymanski LM. Multiple gestations and hypertensive disorders of pregnancy: what do we know? *Curr Hypertens Rep*. 2020;23(1):1.
 50. McDonough CW, Babcock K, Chucru K, et al. Optimizing identification of resistant hypertension: computable phenotype development and validation. *Pharmacoepidemiol Drug Saf*. 2020;29 (11):1393–1401.
 51. Mizuno S, Wagata M, Nagaie S, et al. Development of phenotyping algorithms for hypertensive disorders of pregnancy (HDP) and their application in more than 22,000 pregnant women. *Sci Rep*. 2024;14(1):6292.
 52. Kwan MP. The uncertain geographic context problem. *Ann Assoc Am Geogr*. 2012;102(5):958–968.
 53. USEPA. Technical information about fused air quality surface using downscaling tool: metadata description; 2016. Accessed July 1, 2022. https://www.epa.gov/sites/default/files/2016-07/documents/data_fusion_meta_file_july_2016.pdf
 54. USEPA. Downscaler Model for predicting daily air pollution. July 6, 2015. Accessed June 30, 2022. <https://19january2017snap>

- shot.epa.gov/air-research/downscaler-model-predicting-daily-air-pollution_html
55. van Donkelaar A, Martin RV, Ford B, et al. North American fine particulate matter chemical composition for 0–2022 from satellites, models, and monitors: the changing contribution of wild-fires. *ACS EST Air*. 2024;1(12):1589–1600.
 56. Logue JM, Small MJ, Robinson AL. Evaluating the national air toxics assessment (NATA): comparison of predicted and measured air toxics concentrations, risks, and sources in Pittsburgh, Pennsylvania. *Atmos Environ* (1994). 2011;45(2):476–484.
 57. Daly C, Taylor G, Gibson W. The PRISM approach to mapping precipitation and temperature. In: *10th Conference on Applied Climatology, American Meteorological Society, 0–23 October, Reno NV; 1997*. Accessed May 16, 2022. <http://citereerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.730.5725>
 58. Rothfus LP. *The Heat Index “Equation” (or, More than You Ever Wanted to Know about Heat Index)*. unidata.github.io; 1990. https://unidata.github.io/MetPy/latest/_static/rothfus-1990-heat-index-equation.pdf
 59. Moore RB, McKay LD, Rea AH, et al. User’s guide for the national hydrography dataset plus (NHDPlus) high resolution. US Geological Survey Open-File Report 2019–1096, 66 p., 2019. <https://doi.org/10.3133/ofr20191096>
 60. Rhew IC, Vander Stoep A, Kearney A, Smith NL, Dunbar MD. Validation of the normalized difference vegetation index as a measure of neighborhood greenness. *Ann Epidemiol*. 2011;21(12):946–952.
 61. Thomas J, Zeller L. National walkability index user guide and methodology; May 17, 2021. Accessed April 18, 2022. <https://www.epa.gov/smartgrowth/national-walkability-index-user-guide-and-methodology>
 62. USDA. Introduction to the food access research atlas; 2021. Accessed June 30, 2022. <https://gisportal.ers.usda.gov/portal/apps/experiencebuilder/experience/?id=a53ebd7396cd4ac3a3ed09137676fd40>
 63. Garvin E, Branas C, Keddem S, Sellman J, Cannuscio C. More than just an eyesore: local insights and solutions on vacant land and urban health. *J Urban Health*. 2013;90(3):412–426.
 64. Mennitt D, Sherrill K, Frstrup K. A geospatial model of ambient sound pressure levels in the contiguous United States. *J Acoust Soc Am*. 2014;135(5):2746–2764.
 65. Mennitt DJ, Frstrup KM. Influence factors and spatiotemporal patterns of environmental sound levels in the contiguous United States. *Noise Cont Engng J*. 2016;64(3):342–353.
 66. Elvidge CD, Zhizhin M, Ghosh T, Hsu FC, Taneja J. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sens (Basel)*. 2021;13(5):922.
 67. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health*. 2006;83(6):1041–1062.
 68. Rupasingha A, Goetz SJ, Freshwater D. The production of social capital in US counties. *J Socio Econ*. 2006;35(1):83–101.
 69. Barnett-Ryan C. Introduction to the uniform crime reporting program. *Understanding Crime Stat*. 2007;3:55–89.
 70. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Soft*. 2011;45(3):1–67.
 71. Textor J, Hardt J, Knüppel S. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology*. 2011;22(5):745.
 72. Liu M, Zhang YI, Zhou D. Double/debiased machine learning for logistic partially linear model. *Econom J*. 2021;24(3):559–588.
 73. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. arXiv [cs.LG]. <http://arxiv.org/abs/1603.02754>, March 8, 2016, preprint: not peer reviewed.
 74. Xiang Q, Yuan Y, Song D, et al. Double machine learning to estimate the effects of multiple treatments and their interactions. arXiv [stat.ME]. <http://arxiv.org/abs/2505.12617>, May 18, 2025, preprint: not peer reviewed.
 75. Pedersen M, Stayner L, Slama R, et al. Ambient air pollution and pregnancy-induced hypertensive disorders: a systematic review and meta-analysis. *Hypertension*. 2014;64(3):494–500.
 76. Wylie BJ, Matechi E, Kishashu Y, et al. Placental pathology associated with household air pollution in a cohort of pregnant women from Dar es Salaam, Tanzania. *Environ Health Perspect*. 2017;125(1):134–140.
 77. Stranges S, Navas-Acien A, Rayman MP, Guallar E. Selenium status and cardiometabolic health: state of the evidence. *Nutr Metab Cardiovasc Dis*. 2010;20(10):754–760.
 78. Joseph J. Selenium and cardiometabolic health: inconclusive yet intriguing evidence. *Am J Med Sci*. 2013;346(3):216–220.
 79. Yuan R, Zhang Y, Han J. The association of selenium exposure with the odds of metabolic syndrome: a dose-response meta-analysis. *BMC Endocr Disord*. 2025;25(1):49.
 80. Ekblad MO, Gissler M, Korhonen PE. New theory about the pathophysiology of preeclampsia derived from the paradox of positive effects of maternal smoking. *J Hypertens*. 2022;40(6):1223–1230.
 81. Rodriguez-Lopez M, Escobar MF, Merlo J, Kaufman JS. Reevaluating the protective effect of smoking on preeclampsia risk through the lens of bias. *J Hum Hypertens*. 2023;37(5):338–344.
 82. Giourga C, Papadopoulou SK, Voulgaridou G, Karastogiannidou C, Giaginis C, Pritsa A. Vitamin D deficiency as a risk factor of preeclampsia during pregnancy. *Diseases*. 2023;11(4):158.
 83. AlSubai A, Baqai MH, Agha H, et al. Vitamin D and preeclampsia: a systematic review and meta-analysis. *SAGE Open Med*. 2023;11:20503121231212093.
 84. Liao L, Wei X, Liu M, Gao Y, Yin Y, Zhou R. The association between season and hypertensive disorders in pregnancy: a systematic review and meta-analysis. *Reprod Sci*. 2023;30(3):787–801.
 85. Part C, Le Roux J, Chersich M, et al. Ambient temperature during pregnancy and risk of maternal hypertensive disorders: a time-to-event study in Johannesburg, South Africa. *Environ Res*. 2022;212(Pt D):113596.
 86. Mao Y, Gao Q, Zhang Y, et al. Associations between extreme temperature exposure and hypertensive disorders in pregnancy: a systematic review and meta-analysis. *Hypertens Pregnancy*. 2023;42(1):2288586.
 87. Hastie CE, Mackay DF, Clemens TL, et al. Antenatal exposure to UV-B radiation and preeclampsia: a retrospective cohort study. *J Am Heart Assoc*. 2021;10(13):e020246.
 88. Georgiou M, Morison G, Smith N, Tiegies Z, Chastin S. Mechanisms of impact of blue spaces on human health: a systematic literature review and meta-analysis. *Int J Environ Res Public Health*. 2021;18(5):2486.
 89. Mujahid MS, Diez Roux AV, Morenoff JD, et al. Neighborhood characteristics and hypertension. *Epidemiology*. 2008;19(4):590–598.
 90. Dhingra R, Tamura K, Jayasekera J, Alio AP, Forde AT. A systematic review of the relationship between neighborhood stressors, discrimination, and cardiometabolic outcomes during pregnancy. *NPJ Womens Health*. 2025;3(1):25.
 91. Broadley AJM, Korszun A, Abdelaal E, et al. Inhibition of cortisol production with metyrapone prevents mental stress-induced

- endothelial dysfunction and baroreflex impairment. *J Am Coll Cardiol.* 2005;46(2):344–350.
92. Caplan M, Keenan-Devlin LS, Freedman A, et al. Lifetime psychosocial stress exposure associated with hypertensive disorders of pregnancy. *Am J Perinatol.* 2021;38(13):1412–1419.
93. Shay M, MacKinnon AL, Metcalfe A, et al. Depressed mood and anxiety as risk factors for hypertensive disorders of pregnancy: a systematic review and meta-analysis. *Psychol Med.* 2020;50(13):2128–2140.
94. Bilenko N, Ashin M, Friger M, Fischer L, Sergienko R, Sheiner E. Traffic noise and ambient air pollution are risk factors for pre-eclampsia. *J Clin Med.* 2022;11(15):4552.
95. Chen F, Fu W, Shi O, et al. Impact of exposure to noise on the risk of hypertension: a systematic review and meta-analysis of cohort studies. *Environ Res.* 2021;195:110813.
96. Pedersen M, Halldorsson TI, Olsen SF, et al. Impact of road traffic pollution on pre-eclampsia and pregnancy-induced hypertensive disorders. *Epidemiology.* 2017;28(1):99–106.