An External Exposome-Wide Association Study of Hypertensive Disorders of Pregnancy

Hui Hu[1], **Jinying Zhao**[1], **David A. Savitz**[2], **Mattia Prosperi**[1], **Yi Zheng**[1], **Thomas A. Pearson**[1]

[1]Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, Florida, USA

[2]Department of Epidemiology, School of Public Health, Brown University, Providence, Rhode Island, USA

Abstract

It is widely recognized that exogenous factors play an important role in the development of hypertensive disorders of pregnancy (HDP). However, only a few external environmental factors have been studied, often separately, with no attempt to examine the totality of the external environment, or the external exposome. We conducted an external exposome-wide association study (ExWAS) using the Florida Vital Statistics Birth Records including 819,399 women with live births in 2010–2013. A total of 5,784 factors characterizing women's surrounding natural, built, and social environment during pregnancy from 10 data sources were collected, harmonized, integrated, and spatiotemporally linked to the women based on pregnancy periods using 250m buffers around their geocoded residential addresses. A random 50:50 split divided the data into discovery and replication sets, and a 3-phase procedure was used. In phase 1, associations between HDP and individual factors were examined, and Bonferroni adjustment was performed. In phase 2, an elastic net model was used to perform variable selection among significant variables from phase 1. In phase 3, a multivariable logistic regression model including all variables selected by the elastic net model was fitted. Variables that were significant in both the discovery and replication sets were retained. Among the 528 and 490 variables identified in Phase 1, 232 and 224 were selected by the elastic net model in Phase 2, and 67 and 48 variables remained statistically significant in Phase 3 in the discovery and replication sets, respectively. A total of 12 variables were significant in both the discovery and replication sets, including air toxicants (e.g., 2,2,4-trimethylpentane), meteorological factors (e.g., omega or vertical velocity at 125mb pressure level), neighborhood crime and safety (e.g., burglary rate), and neighborhood sociodemographic status (e.g., urbanization). This is the first large external exposome study of HDP. It confirmed some of the previously reported associations and generated unexpected predictors within the environment that may warrant more focused evaluation.

Address correspondence to Hui Hu, University of Florida, College of Public Health and Health Professions and College of Medicine, Department of Epidemiology, 2004 Mowry Road, CTRB 4224, Gainesville, FL 32610 USA. Telephone: 352-294-5944. Fax: 352-273-5365. huihu@ufl.edu.

Keywords

Exposome; Hypertensive Disorders of Pregnancy; Preeclampsia

Introduction

Hypertensive disorders of pregnancy (HDP) are among the most common medical problems encountered during pregnancy, affecting up to 10% of all pregnancies. These medical problems are characterized by chronic hypertension or high blood pressure that develops after 20 weeks of gestation because blood volume change during pregnancy leads to higher demand on the cardiovascular system (American College of Obstetricians abd Gynecologists 2013; Yoder et al. 2009). HDP are important risk factors for increased neonatal and maternal morbidity and mortality (Duley 2009; Lo et al. 2013). Maternal HDP are associated with higher risks of small for gestational age, preterm delivery, low birthweight, and hospitalization for a wide range of neonatal diseases (Pinheiro et al. 2016). Among mothers, HDP is predictive of pitting edema, endothelial abnormalities, liver and renal dysfunction, and increased risk of cardiovascular disease, stroke and Type II diabetes later in life (Bauer and Cleary 2009; Bellamy et al. 2007; Brown et al. 2013; Duley 2009; Garovic et al. 2010; Lykke et al. 2009; McDonald et al. 2008; Tooher et al. 2016; Wang et al. 2012; Wu et al. 2017).

Similar to many other cardiovascular outcomes, HDP has large geographic disparities (Gebreab et al. 2015; Hu et al. 2019; Organization 1988; Shoff 2012), suggesting that environment may play an important role in its development. While previous studies have identified multiple exposures associated with HDP in the natural, built, and social environments (Clausen et al. 2006; Grazuleviciene et al. 2014; Hu et al. 2014; Messer et al. 2012; Morales et al. 2016; Osorio-Yanez et al. 2016; Strand et al. 2011; Tran et al. 2015; Vinikoor-Imler et al. 2012), these studies only focused on a subset of preselected environmental factors, and predominantly assessed these factors in isolation without considering the totality of environment, or the exposome (Wild 2012).

The exposome includes internal (e.g. metabolism), specific external (e.g. pollutants), and general external factors (e.g. social capital) (Wild 2012). In this study, we focused on the specific and general external environmental factors as the ones most plausibly amenable to modification. We conducted an external exposome-wide association study (ExWAS) of HDP using the Florida Vital Statistics Birth Records. The term ExWAS is increasingly used given its more comprehensive nature compared with the environment-wide association study (Juarez and Matthews-Juarez 2018; Nieuwenhuijsen et al. 2019), or EWAS, which was first introduced by Patel et al. (2010), borrowing the idea from genome-wide association studies (GWAS) that identify genetic factors associated with diseases. Unlike genetic factors which are stable and unmodifiable, environmental factors have large spatiotemporal variability and can be modified at multiple levels such as behavior changes at the individual level and policy changes at the community level. Therefore, putative environmental factors identified by ExWAS can be used not only for understanding disease risk factors, but potentially for disease prevention or intervention as well. In this study, we analyzed Florida Vital Statistics

Birth Records to examine the association between the external exposome and risk of HDP by integrating data on a wide range of environmental factors characterizing women's surrounding natural, built, and social environment. Using the agnostic and hypothesis-free ExWAS approach with integration of multi-source environmental big data, we aimed to identify novel environmental factors associated with HDP.

Materials and Methods

Study population

Individual records of all registered live births in Florida between 2010 and 2014 (n=1,064,604) were obtained from the Florida Department of Health (Jacksonville, Florida, http://www.floridahealth.gov/certificates/certificates/). Births with maternal residential addresses unsuccessfully geocoded (n=704) or outside Florida (n=5,292) were excluded. To avoid fixed cohort bias (Barnett 2011), we included women based on estimated conception date instead of delivery date, and a total of 846,579 records with conception date between January 1, 2010 and December 31, 2013 were included. For women with multiple births, only one record was randomly chosen and retained for each pair and duplicate records were excluded (n=13,232). We then excluded women with pre-pregnancy hypertension (n=12,178) or missing information on HDP (n=1,770), and 819,399 women were included in the analyses. The research protocol for this study was approved by the Institutional Review Boards at the University of Florida (IRB201701782) and the Florida Department of Health (2017–43-UFL).

Outcome assessment

Diagnoses of pre-pregnancy hypertension, gestational hypertension or preeclampsia, and eclampsia obtained from the Florida Vital Statistics Birth Record data. Similar to previous environmental studies on HDP (Hu et al. 2014), HDP included gestational hypertension, preeclampsia, or eclampsia. HDP was assessed aggregately since the Florida Vital Statistics Birth Record data combined gestational hypertension and preeclampsia together.

External Exposome Measures

Data on a variety of natural, built, and social environment measures were obtained from ten different sources and spatiotemporally linked to each woman to assess their external exposome during pregnancy. Several different temporal exposure windows were used in the linkage, depending on the type and temporal scale of environment measures. These windows include the first trimester, the second trimester, and the first and second trimesters. For environment measures with a temporal scale less than 1-year, all three windows were used in the linkage. For those with a temporal scale greater than or equal to 1-year, only the first and second trimesters was used in the linkage. All measures were spatially linked to each woman using a 250m buffer around each woman's geocoded residential address at delivery based on area-weighted averages. Table 1 shows a summary of the external exposome measures. A total of 5,784 variables covering 10 categories were finally included.

Natural environment—Air toxicant measures were generated using data from the National Air Toxic Assessment (NATA). Concentration estimates of 119 air toxicants are

available at the census tract level for 2005, 2011, and 2014. Linear interpolations were performed to construct measures in 2010 and 2012–2013. We excluded 30 air toxicants with missing data exceeding 5%, and a total of 89 air toxicants were included. In addition, data on fine particulate matter and ozone were obtained from the US Environmental Protection Agency (EPA) and Center for Disease Control and Prevention (CDC)'s National Environmental Public Health Tracking Network (U.S. EPA 2014). A Bayesian space-time downscaler model is used to fuse daily $PM_{2.5}$ and O_3 monitoring data from the National Air Monitoring Stations/State and Local Air Monitoring Stations (NAMS/SLAMS) with 12 km gridded output from the Models-3/Community Multiscale Air Quality (CMAQ) model (EPA 2016). Daily estimates were obtained at census tract level for 2010–2014. Furthermore, we obtained meteorological measures from the National Centers for Environmental Prediction's North American Regional Reanalysis (NARR) (Environmental Modeling Center 1979). The NARR provides a high-resolution meteorology database with a daily temporal resolution and a 32km spatial resolution. A total of 279 meteorological measures were obtained. We further excluded 20 measures with missing data exceeding 5% and 2 measures with only one unique value (i.e. all of the values for the measure were the same). A total of 257 meteorological measures were included.

Built environment—Food access measures were obtained from USDA's Food Access Research Atlas (USDA 2014). The data were available in 2010 and 2015, and linear interpolation was conducted to construct measures in 2011–2014. A total of 44 measures were included. Green space was measured using the Normalized Difference Vegetation Index (NDVI) from the NASA's TERRA/MODIS, which has been shown to be effective in measuring neighborhood greenness for epidemiological studies (Rhew et al. 2011). The data have a 16-day temporal resolution and a 250-meter spatial resolution. Walkability information were obtained from the Walk Score API, which measures walkability on a scale from 0–100 based on walking routes to destinations including grocery stores, schools, parks, restaurants, and retail stores (Brewster et al. 2009). The data are available with a 0.0015 decimal degrees spatial resolution.

Social environment—Sociodemographic status and housing measures were generated using data from the American Community Survey (ACS) 5-year estimates at the 2010 census block group level. Since block group level estimates were only available from the 5-year estimates data, we used the middle-year of the 5-year period as the index year in the linkage. A total of 3,459 variables that are available during 2008–2012 and 2012–2016 ACS were obtained using the "totalcensus" package in R, and we included 3,242 variables with nonmissing values. All the included variables were count variables which we used to calculate proportions. A total of 5,650 proportion variables were generated. We assigned 0 to proportions where the denominator is 0. We excluded 24 measures with only 1 unique value. Pairwise Pearson correlations were computed, and we further excluded 696 measures with absolute correlations>0.99 with another measure. A total of 4,930 measures were included. Deprivation indices such as the neighborhood deprivation index (Messer et al. 2006) and the area deprivation index (University of Wisconsin School of Medicine and Public Health 2015) were not included since we aimed to examine the association between each social environmental factor and HDP. Address vacancy measures at the 2010 census-tract level in

2010–2014 were obtained from the US Department of Housing and Urban Development (HUD) aggregated USPS administrative data. A total of 122 measures were included. In addition, ten social capital measures were constructed using the Census Business Pattern data based on the North American Industry Classification System (NACIS) codes (Rupasingha et al. 2006) at the 2010 census 5-digit ZIP code tabulation area (ZCTA5) level. Furthermore, seven county-level annual crime measures (i.e. forcible sex offenses, murder, robbery, aggravated assault, burglary, larceny, and motor vehicle theft rate) were obtained from the Uniform Crime Reporting Program in 2010–2014 (Barnett-Ryan 2007).

Covariates

Women's age at delivery was categorized into six groups, with 5-year increments for women aged 20–40 years old and two additional groups for <20 and 40 years old. Race/ethnicity was categorized as non-Hispanic White, non-Hispanic Black, Mexican American, Puerto Rican, Cuban American, Haitian American, and others. In addition, education (<high school, high school or equivalent, some college, college graduate, or >college), marital status (not currently married or currently married), and status of the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC, which is a program for low-income families in the US) were also obtained. Pregnancy smoking status was categorized into two levels: non-smokers or smokers. Pre-pregnancy body mass index (BMI) was categorized into four levels: underweight (<18.5), normal (18.5–24.9), overweight (25.0–29.9), and obese (30.0). Parity (nulliparous or parous), season [warm (June-November) or cool (December-May)] and year (2010, 2011, 2012, or 2013) of conception were also included.

Statistical analysis

Normalization transformations of all continuous environment variables were performed using the bestNormalize package in R (Peterson 2018), which implements several transformation methods including the log, square root, exponential, arcsinh, Box Cox, and Yeo-Johnson transformations. The best transformation was determined based on the Pearson P statistics. Supplemental Tables 1 and 2 shows the chosen transformations and parameters for all the environment variables. All continuous variables were also z-score standardized (mean=0 and standard deviation=1). Pairwise Pearson correlation coefficients were calculated. For every pair of perfectly correlated variables (absolute value of correlation coefficients >0.99), we excluded one of the variables. A total of 443 variables were excluded, and 5,784 variables were included in the external ExWAS. Supplemental Table 3 shows the excluded variables and corresponding included variables.

Missing data for all external exposome factors and covariates were imputed using the chained equations method by the mice package in R. We did not impute HDP since previous study showed that there is no gain in power from imputing the outcome variable (White et al. 2011). A variables was considered as a predictor in the imputation model if its proportion of nonmissing values among women with missing values in the variable to be imputed was larger than 40% and they were correlated (i.e. with the absolute correlation value>0.4) with the variable to be imputed or the probability of the variable being missing. HDP was forced in as a predictor in the imputation model, and we allowed the number of predictors in the imputation model to be 25 at maximum (Van Buuren 2018). We imputed a single dataset

given the minimal impacts of the imputation procedure due to the large sample size and small fractions of missing data.

The analyses procedure in this external ExWAS was designed based on the results from a simulation study by Agier et al. (2016) which compared the performance of several statistical methods in assessing exposome-health associations, including the EWAS approach (Patel et al. 2010), EWAS followed by a multivariable regression step including the identified hits (EWAS-MLR) (Patel et al. 2013), elastic net (Zou and Hastie 2005), sparse partial least squares (sPLS) regression (Chun and Kele 2010), the Graphical Unit Evolutionary Stochastic Search (GUESS) algorithm (Bottolo et al. 2013), and the deletion/substitution/addition (DSA) sequential algorithm (Sinisi and Van Der Laan 2004). While GUESS and DSA provided marginally better balance between sensitivity and false discovery proportions (FDP), they are more compute-intensive and are not feasible to be applied in the current study with a total of 5,784 exposome factors among 819,399 women (Agier et al. 2016). Instead, a three-phase procedure was used to maximize the strengths of the EWAS, elastic net, and EWAS-MLR approaches. We randomly split the data into discovery (50%) and replication (50%) sets and conducted the same analyses in Phases 1 to 3 separately in the two sets. In Phase 1, we used the EWAS method because it has been shown to have the highest sensitivity among all the methods compared (Agier et al. 2016; Shoff 2012). Specifically, we considered all the 5,784 environment variables for associations with HDP after accounting for multiple comparisons. Logistic regression models were fitted for each environment measure after adjusting for all the covariates. To account for the multiple testing, we set the significance threshold as 0.86×10^{-7} using Bonferroni adjustment (Dunn 1961). To account for the high FDP of the EWAS method (Agier et al. 2016), in Phase 2, elastic net model was used for variable selections among significant variables from Phase 1. All covariates included in Phase 1 were also included in the elastic net model. Elastic net model has been shown to have substantially low FDP compared with the EWAS method (Agier et al. 2016). Hyper-parameters of the elastic net model were tuned based on 5-fold cross-validated AUC. In Phase 3, we used the EWAS-MLR approach to estimate the effect sizes by fitting a multivariable logistic regression model including all selected variables from Phase 2 as well as all the covariates. Previous simulation study has shown that the EWAS-MLR approach performs well in estimating the true predictor coefficient values (Agier et al. 2016). Odds ratios (ORs) and 95% confidence intervals (95% CIs) were reported. Variables remained significant in Phase 3 from both the discovery and replication sets are considered as validated. We further generated a correlation heatmap showing the pairwise Pearson correlations of the validated variables. Figure 1 shows the flow chart summarizing the external ExWAS methods employed in this study. To examine the robustness of the results, sensitivity analyses were conducted by additionally adjusting for insurance (i.e. Medicaid, private, or self-pay) and prenatal care within the first 2 trimesters (i.e. yes or no) in all models. All analyses were performed using R statistical software (version 3.6; R Development Core Team).

Results

Of the 819,399 women included in this study, 42,746 (5.2%) had HDP identified in the birth record. Table 2 shows the distribution of women's characteristics by HDP status. Women

with HDP were older and more likely to be non-Hispanic Black, unmarried, WIC recipient, parous, and have higher BMI, confirming known predictors.

Figure 2 shows the volcano plot summarizing results from Phase 1. After accounting for multiple comparisons using the Bonferroni adjustment, a total of 528 and 490 variables were significantly associated with HDP in the discovery and replication sets, respectively. Among them, 374 variables were significant in both the discovery and replication sets. Supplemental Tables 1 and 2 shows the ORs, 95% CIs, and p-values for each of the 5,784 environment variables from Phase 1.

Among the 528 and 490 significant variables in Phase 1, a total of 232 and 224 were further identified by the elastic net model in Phase 2 from the discovery and replication sets, respectively. Supplemental Tables 4 and 5 show the estimated coefficients from the elastic net model.

In Phase 3, all variables identified in Phase 2 were simultaneously included in a multivariable logistic regression model after adjusting for the covariates, and 67 and 48 variables remained statistically significant in the discovery and replication sets, respectively (Supplemental Tables 4 and 5). A total of 12 variables were significant in both the discovery and replication sets and were considered as validated.

Figure 3 shows the pairwise correlations of the 12 validated variables. The strongest correlation was observed between omega at 125mb pressure level during the first two trimesters and omega at 100mb pressure level during the second trimester, with a Pearson correlation coefficient of 0.67. All other correlation coefficients have absolute values below 0.5.

Table 3 shows the adjusted ORs and 95% CIs (for each standard deviation increase) for these variables. Six of the 12 variables were air toxicants, including exposure to 2,2,4-trimethylpentane ($OR_{Discovery}$: 1.15, 95% CI: 1.08, 1.23), acrolein ($OR_{Discovery}$: 1.09, 95% CI: 1.05, 1.13), 1,1,1-tricholorethane ($OR_{Discovery}$: 1.06, 95% CI: 1.01, 1.11), maleic anhydride ($OR_{Discovery}$: 1.04, 95% CI: 1.01, 1.06), propylene oxide ($OR_{Discovery}$: 1.02, 95% CI: 1.01, 1.04), and methyl tert-butyl ether ($OR_{Discovery}$: 0.90, 95% CI: 0.85, 0.94) during the first two trimesters. Two were crime and safety measures, including burglary rate ($OR_{Discovery}$: 1.07, 95% CI: 1.05, 1.09) and forcible sex offense rate ($OR_{Discovery}$: 1.03, 95% CI: 1.01, 1.06) during the first two trimesters. Two were sociodemographic variables, including proportion of individuals living in a census defined place (i.e. a concentration of population which has a name such as a city, town, or village) among workers 16 years and over ($OR_{Discovery}$: 1.05, 95% CI: 1.01, 1.10), and proportion of individuals who speak English very well among those aged 18–64 years who speak Spanish ($OR_{Discovery}$: 0.98, 95% CI: 0.96, 0.99) during the first two trimesters. The other two were meteorology variables, including omega at 125mb pressure level during the first two trimesters ($OR_{Discovery}$: 1.13, 95% CI: 1.04, 1.23) and omega at 100mb pressure level during the second trimester ($OR_{Discovery}$: 0.91, 95% CI: 0.87, 0.96).

Consistent results were observed in the sensitivity analyses which additionally adjusted for insurance and prenatal care in all models. Supplemental Table 6 shows the number of

variables retained in each phase. The same 12 validated variables were found in the sensitivity analyses. Supplemental Table 7 shows the ORs and 95% CIs of the 12 validated variables in Phase 3 from the original analyses and the sensitivity analyses.

Discussion

To our knowledge, this is the largest external ExWAS that has been conducted to date with numerous factors in the natural, built, and social environment included. This is also the first study to agnostically and comprehensively assess and identify external exposome factors associated with HDP. We assessed the association of 5,784 external exposome variables with HDP using a multi-phase procedure. After accounting for multiple testing and high correlations among the exposures, 12 variables characterizing the natural (i.e., air toxicants and meteorology)and social environment (i.e., sociodemographic status and housing, and crime and safety) were identified to be significantly associated with HDP.

Previous studies have found that exposure to air pollution during pregnancy may be associated with HDP (Hu et al. 2014). However, most studies focused on criteria pollutants, and little is known on other air toxicants. We identified 6 novel air toxicants that were associated with HDP. Five of them were positively associated with HDP, which were mainly presented in gasoline (i.e. 2,2,4-trimethylpentane and acrolein), used as solvent to dissolve other substances (i.e. 1,1,1-trichloroethane), or for the production of polymers (i.e. maleic anhydride and propylene oxide) (National Toxicology Program 2019). While the biological mechanisms underlying these associations are largely unknown, they may share the same potential pathways for other air pollutants such as $PM_{2.5}$ (Hu et al. 2020), including inflammation, oxidative stress, endothelial dysfunction, and DNA methylation (Abraham et al. 2018). One air toxicant that was negatively associated with HDP, methyl tert-butyl ether, was also present in tobacco(National Toxicology Program 2019), suggesting that it may share some common mechanisms underlying the reduced risk of HDP observed among women with tobacco use during pregnancy (Wikström et al. 2010).

Meteorology has been linked to the risk of HDP (Beltran et al. 2014; Tran et al. 2015). However, previous studies predominantly focused on temperature, humidity, and precipitation (Beltran et al. 2014), and scarce evidence is available for other meteorological factors. In this external ExWAS, in addition to these widely studied factors, we examined many other meteorological factors. We found that vertical velocity measured by the omega equation was associated with HDP. Most high impact weather occurs when vertical velocity increases (Dostalek et al. 2017). The observed positive association between vertical velocity during the first two trimesters and HDP is consistent with previous studies. Yackerson et al. (2007) found that number of days with strong wind is associated with preeclampsia, and Haelterman et al. (2007) found that exposure to extreme weather in the first 20 weeks of pregnancy is associated with increased risk of HDP. We also observed a negative association between vertical velocity during the second trimester and HDP in the multivariable models, suggesting that early pregnancy might be a critical window of exposure to extreme weather. Previous studies showed mixed findings on the association between temperature, humidity, and precipitation (Beltran et al. 2014). In this external ExWAS, precipitation is not

significantly associated with HDP, and temperature and humidity are only significantly associated with HDP in Phase 1.

Built environment and HDP have also been examined in previous studies. Pregnancy exposure to an environment with better food quality and quantity (Morales et al. 2016), sidewalks/trails (Messer et al. 2012), and parks/recreational services (Grazuleviciene et al. 2014) has been associated with reduced risks of HDP, while no association was found between green space and HDP (Laurent et al. 2013; Young et al. 2016). In this external ExWAS, we found that social capital, food access, walkability, and green space were significantly associated with HDP in only the discovery set or the replication set, and no significant associations were found in both the discovery and replication sets.

Living in a neighborhood with disadvantaged social environment has been associated with hypertension and cardiovascular diseases in the general population (Agyemang et al. 2007; Cozier et al. 2007; Cubbin et al. 2000; Cubbin et al. 2006; Morenoff et al. 2007; Mujahid et al. 2008; Roux et al. 2001). However, only a few studies have investigated the relationship between social environment and HDP and conflicting results have been observed (Agyemang et al. 2009; Clausen et al. 2006; Gudmundsson et al. 1997; Messer et al. 2012; Vinikoor-Imler et al. 2012). In this external ExWAS, we found that urbanization and acculturation are associated with HDP. The observed positive association between proportion of individuals living in a census defined place (i.e. a concentration of population which has a name such as a city, town, or village) among workers 16 years and over and risk of HDP is consistent with the elevated risk of HDP observed in urban areas (Callaway et al. 2006; van Middendorp et al. 2013). In addition, women living in neighborhoods with higher proportion of individual who speak English well among those who speak Spanish had lower risk of HDP. Furthermore, burglary and forcible sex offenses rates were associated with increased risk of HDP, which is consistent with a previous study in Chicago (Mayne et al. 2018).

Similar to GWAS, many of the statistically significant associations observed in this external ExWAS have small effect sizes. There are two main contributors to the small effect sizes in the context of ExWAS: 1) the potential exposure misclassifications, which may bias the estimates toward the null, and 2) the high correlations among the exposures, which lead to attenuated effect sizes in multivariate models. Such small effect sizes are well below the usual levels that are informative for etiologic studies. However, it is possible that some exposures may share common or similar etiologic mechanisms. Populations are usually exposed to a mixture of these exposures, and the additive contributions from exposures with small effect sizes may be large.

Our study has several strengths. By using the agnostic ExWAS approach, we included a variety of external exposome measures, which addressed limitations of previous studies that assessed only a small fraction of these factors separately. To the best of our knowledge, only three external ExWAS have been conducted (Lynch et al. 2017; Mooney et al. 2017; Nieuwenhuijsen et al. 2019). Two of them used cross-sectional designs focusing on prostate cancer and physical activity (Lynch et al. 2017; Mooney et al. 2017), and only Nieuwenhuijsen et al. (2019) considered temporal information when assessing external

exposome factors associated with birth weight. In addition, none of these studies simultaneously considered natural, built, and social environment. In this external ExWAS, we were able to spatiotemporally link data characterizing natural, built, and social environment to each woman to account for the temporal dynamic nature of environment exposures. Furthermore, by using the Florida Vital Statistics Birth Records, we were able to obtain an extremely large sample size with minimal selection bias and multiple individual-level confounders adjusted. Finally, the findings from this study also suggest that environmental factors have the potential to be used to identify women at high risk of HDP, which can guide targeted early pregnancy interventions such as administrations of aspirin (Rolnik et al. 2017).

Several limitations also need to be noted. The accuracy and detail on HDP based on the birth certificate alone is known to be subject to misclassification and precludes examination of the subsets of hypertensive disorders that may have differing etiologies (Lydon-Rochelle et al. 2005; Milic et al. 2018; Stuart 2018). Previous studies showed that HDP determined by the birth certificate has very high specificity (>0.99) (Dietz et al. 2015; DiGiuseppe et al. 2002; Lydon-Rochelle et al. 2005), suggesting that the misclassification may not cause substantial bias in estimating the associations (Rothman et al. 2008). We treated HDP as an aggregate outcome. However, it is recognized that HDP with different subtype, onset, and severity have distinct pathogenic mechanisms as well as maternal and fetal complications (Valensise et al. 2008). Therefore, although they may share many risk and protective factors in common, specific risk and protective factors may exist. Studies in the future are warranted to consider the subtype, onset, and severity of HDP. We did not consider nonlinear association and potential interactions, lacked information on daily mobility and residential history during pregnancy, and many of the exposures considered are subject to potential measurement error. In addition, although many environment factors have been included to characterize the external exposome, this list is not exhaustive, and continuing efforts are needed to further improve measurement of the external exposome. Furthermore, only internal validation was performed.

Conclusion

This external ExWAS provides new insights into the role of the external exposome in HDP and has implications for precision health studies by incorporating information beyond omics. We confirmed some of the previously reported associations (i.e. crime rate, and urbanization), and identified novel factors associated with HDP, including vertical velocity (assessed by the omega equation), neighborhood acculturation, and air toxicants such as 2,2,4-trimethylpentane, acrolein, 1,1,1-trichloroethane, maleic anhydride, propylene oxide, and methyl tert-butyl ether, and meteorological factor. Future studies are also warranted to confirm the novel associations identified by this study with improved measurements and to better understand the mechanisms underlying these associations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

ACS	American Community Survey
BMI	Body Mass Index
CDC	Center for Disease Control and Prevention
EPA	Environmental Protection Agency
ExWAS	Exposome-wide Association Study
EWAS	Environment-wide Association Study
GWAS	Genome-wide Association Study
HDP	Hypertensive Disorders of Pregnancy
HUD	Department of Housing and Urban Development
NACIS	North American Industry Classification System
NDVI	Normalized Difference Vegetation Index
WIC	The Special Supplemental Nutrition Program for Women, Infants and Children

References

Abraham E, Rousseaux S, Agier L, Giorgis-Allemand L, Tost J, Galineau J, et al. 2018 Pregnancy exposure to atmospheric pollution and meteorological conditions and placental DNA methylation. Environment international 118:334–347. [PubMed: 29935799]

Agier L, Portengen L, Chadeau-Hyam M, Basagaña X, Giorgis-Allemand L, Siroux V, et al. 2016 A systematic comparison of linear regression–based statistical methods to assess exposome-health associations. Environmental health perspectives 124:1848–1856. [PubMed: 27219331]

Agyemang C, van Hooijdonk C, Wendel-Vos W, Ujcic-Voortman JK, Lindeman E, Stronks K, et al. 2007 Ethnic differences in the effect of environmental stressors on blood pressure and hypertension in the netherlands. BMC Public Health 7:1. [PubMed: 17199891]

Agyemang C, Vrijkotte T, Droomers M, Van der Wal M, Bonsel G, Stronks K. 2009 The effect of neighbourhood income and deprivation on pregnancy outcomes in amsterdam, the netherlands. Journal of epidemiology and community health 63:755–760. [PubMed: 19679715]

American College of Obstetricians abd Gynecologists. 2013 Hypertension in pregnancy. Report of the american college of obstetricians and gynecologists' task force on hypertension in pregnancy. Obstetrics and gynecology 122:1122. [PubMed: 24150027]

Barnett-Ryan C 2007 Introduction to the uniform crime reporting program. 2007), Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR:55–89.

Barnett AG. 2011 Time-dependent exposures and the fixed-cohort bias. Environmental health perspectives 119:a422.

Bauer ST, Cleary KL. 2009 Cardiopulmonary complications of pre-eclampsia. Seminars in perinatology 33:158–165. [PubMed: 19464506]

Bellamy L, Casas JP, Hingorani AD, Williams DJ. 2007 Pre-eclampsia and risk of cardiovascular disease and cancer in later life: Systematic review and meta-analysis. Bmj 335:974. [PubMed: 17975258]

Beltran A, Wu J, Laurent O. 2014 Associations of meteorology with adverse pregnancy outcomes: A systematic review of preeclampsia, preterm birth and birth weight. International journal of environmental research and public health 11:91–172.

Bottolo L, Chadeau-Hyam M, Hastie DI, Zeller T, Liquet B, Newcombe P, et al. 2013 Guess-ing polygenic associations with multiple phenotypes using a gpu-based evolutionary stochastic search algorithm. PLoS genetics 9:e1003657. [PubMed: 23950726]

Brewster M, Hurtado D, Olson S, Yen J. 2009 Walkscore. Com: A new methodology to explore associations between neighborhood resources, race, and health. APHA Poster Presentation Retrieved Electronically From: https://aphaconfex.com/recording/apha/137am/pdf/free/4db77adf5df9fff0d3caf5cafe28f496/paper205082_1pdf.

Brown MC, Best KE, Pearce MS, Waugh J, Robson SC, Bell R. 2013 Cardiovascular disease risk in women with pre-eclampsia: Systematic review and meta-analysis. European journal of epidemiology 28:1–19. [PubMed: 23397514]

Callaway LK, Chang AM, McIntyre HD, Prins JB. 2006 The prevalence and impact of overweight and obesity in an australian obstetric population. Medical Journal of Australia 184:56–59. [PubMed: 16411868]

Chun H, Keleş S. 2010 Sparse partial least squares regression for simultaneous dimension reduction and variable selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72:3–25. [PubMed: 20107611]

Clausen T, Øyen N, Henriksen T. 2006 Pregnancy complications by overweight and residential area. A prospective study of an urban norwegian cohort. Acta obstetricia et gynecologica Scandinavica 85:526–533. [PubMed: 16752229]

Cozier YC, Palmer JR, Horton NJ, Fredman L, Wise LA, Rosenberg L. 2007 Relation between neighborhood median housing value and hypertension risk among black women in the united states. American journal of public health 97:718–724. [PubMed: 17329664]

Cubbin C, Hadden WC, Winkleby MA. 2000 Neighborhood context and cardiovascular disease risk factors: The contribution of material deprivation. Ethnicity & disease 11:687–700.

Cubbin C, Sundquist K, Ahlén H, Johansson S-E, Winkleby MA, Sundquist J. 2006 Neighborhood deprivation and cardiovascular disease risk factors: Protective and harmful effects. Scandinavian journal of public health 34:228–237. [PubMed: 16754580]

Dietz P, Bombard J, Mulready-Ward C, Gauthier J, Sackoff J, Brozicevic P, et al. 2015 Validation of selected items on the 2003 us standard certificate of live birth: New york city and vermont. Public Health Reports 130:60–70. [PubMed: 25552756]

DiGiuseppe DL, Aron DC, Ranbom L, Harper DL, Rosenthal GE. 2002 Reliability of birth certificate data: A multi-hospital comparison to medical records information. Maternal and child health journal 6:169–179. [PubMed: 12236664]

Dostalek JF, Schubert WH, DeMaria M. 2017 Derivation and solution of the omega equation associated with a balance theory on the sphere. Journal of Advances in Modeling Earth Systems 9:3045–3068.

Duley L 2009 The global impact of pre-eclampsia and eclampsia. Seminars in perinatology 33:130–137. [PubMed: 19464502]

Dunn OJ. 1961 Multiple comparisons among means. Journal of the American statistical association 56:52–64.

Environmental Modeling Center. 1979 National centers for environmental prediction. National Weather Service/NOAA/US Department of Commerce, NCEP Climate Forecast System Reanalysis (CFSR) Selected Hourly Time-Series Products.

EPA. 2016 Downscaler model for predicting daily air pollution. Available: https://www.epa.gov/air-research/downscaler-model-predicting-daily-air-pollution [accessed 2019/1/10].

Garovic VD, Bailey KR, Boerwinkle E, Hunt SC, Weder AB, Curb D, et al. 2010 Hypertension in pregnancy as a risk factor for cardiovascular disease later in life. Journal of hypertension 28:826. [PubMed: 20087214]

Gebreab SY, Davis SK, Symanzik J, Mensah GA, Gibbons GH, Diez-Roux AV. 2015 Geographic variations in cardiovascular health in the united states: Contributions of state-and individual-level factors. Journal of the American Heart Association 4:e001673. [PubMed: 26019131]

Grazuleviciene R, Dedele A, Danileviciute A, Vencloviene J, Grazulevicius T, Andrusaityte S, et al. 2014 The influence of proximity to city parks on blood pressure in early pregnancy. Int J Environ Res Public Health 11:2958–2972. [PubMed: 24619158]

Gudmundsson S, Björgvinsdóttir L, Molin J, Gunnarsson G, Marsal K. 1997 Socioeconomic status and perinatal outcome according to residence area in the city of malmö. Acta obstetricia et gynecologica Scandinavica 76:318–323. [PubMed: 9174424]

Haelterman E, Marcoux S, Croteau A, Dramaix M. 2007 Population-based study on occupational risk factors for preeclampsia and gestational hypertension. Scandinavian journal of work, environment & health:304–317.

Hu H, Ha S, Roth J, Kearney GD, Talbott EO, Xu X. 2014 Ambient air pollution and hypertensive disorders of pregnancy: A systematic review and meta-analysis. Atmos Environ (1994) 97:336–345.

Hu H, Xiao H, Zheng Y, Yu BB. 2019 A bayesian spatio-temporal analysis on racial disparities in hypertensive disorders of pregnancy in florida, 2005–2014. Spatial and spatio-temporal epidemiology 29:43–50. [PubMed: 31128630]

Hu H, Bian J, Zhao J. 2020 Ambient air pollution and preeclampsia: Looking back and moving forward. Hypertension (Dallas, Tex: 1979):HYPERTENSIONAHA11913269.

Juarez PD, Matthews-Juarez P. 2018 Applying an exposome-wide (exwas) approach to cancer research. Frontiers in oncology 8.

Laurent O, Wu J, Li L, Milesi C. 2013 Green spaces and pregnancy outcomes in southern california. Health & place 24:190–195. [PubMed: 24140704]

Lo JO, Mission JF, Caughey AB. 2013 Hypertensive disease of pregnancy and maternal mortality. Current opinion in obstetrics & gynecology 25:124–132. [PubMed: 23403779]

Lydon-Rochelle MT, Holt VL, Cárdenas V, Nelson JC, Easterling TR, Gardella C, et al. 2005 The reporting of pre-existing maternal medical conditions and complications of pregnancy on birth certificates and in hospital discharge data. American journal of obstetrics and gynecology 193:125–134. [PubMed: 16021070]

Lykke JA, Langhoff-Roos J, Sibai BM, Funai EF, Triche EW, Paidas MJ. 2009 Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother. Hypertension 53:944–951. [PubMed: 19433776]

Lynch SM, Mitra N, Ross M, Newcomb C, Dailey K, Jackson T, et al. 2017 A neighborhood-wide association study (nwas): Example of prostate cancer aggressiveness. PloS one 12:e0174548. [PubMed: 28346484]

Mayne SL, Pool LR, Grobman WA, Kershaw KN. 2018 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 72:230–236. [PubMed: 29305526]

McDonald SD, Malinowski A, Zhou Q, Yusuf S, Devereaux PJ. 2008 Cardiovascular sequelae of preeclampsia/eclampsia: A systematic review and meta-analyses. American heart journal 156:918–930. [PubMed: 19061708]

Messer LC, Laraia BA, Kaufman JS, Eyster J, Holzman C, Culhane J, et al. 2006 The development of a standardized neighborhood deprivation index. 83:1041–1062.

Messer LC, Vinikoor-Imler LC, Laraia BA. 2012 Conceptualizing neighborhood space: Consistency and variation of associations for neighborhood factors and pregnancy health across multiple neighborhood units. Health Place 18:805–813. [PubMed: 22551891]

Milic NM, Codsi E, Tobah YSB, White WM, Kattah AG, Weissgerber TL, et al. Electronic algorithm is superior to hospital discharge codes for diagnoses of hypertensive disorders of pregnancy in historical cohorts In: Proceedings of the Mayo Clinic Proceedings, 2018, Vol. 93Elsevier, 1707–1719.

Mooney SJ, Joshi S, Cerdá M, Kennedy GJ, Beard JR, Rundle AG. 2017 Contextual correlates of physical activity among older adults: A neighborhood environment-wide association study (ne-was).AACR.

Morales ME, Epstein MH, Marable DE, Oo SA, Berkowitz SA. 2016 Peer reviewed: Food insecurity and cardiovascular health in pregnant women: Results from the food for families program, chelsea, massachusetts, 2013–2015. Preventing Chronic Disease 13.

Morenoff JD, House JS, Hansen BB, Williams DR, Kaplan GA, Hunte HE. 2007 Understanding social disparities in hypertension prevalence, awareness, treatment, and control: The role of neighborhood context. Social science & medicine 65:1853–1866. [PubMed: 17640788]

Mujahid MS, Roux AVD, Morenoff JD, Raghunathan TE, Cooper RS, Ni H, et al. 2008 Neighborhood characteristics and hypertension. Epidemiology 19:590–598. [PubMed: 18480733]

National Toxicology Program. 2019 Report on carcinogens.Department of Health and Human Services.

Nieuwenhuijsen MJ, Agier L, Basagaña X, Urquiza J, Tamayo-Uria I, Giorgis-Allemand L, et al. 2019 Influence of the urban exposome on birth weight. Environmental health perspectives 127:047007.

Organization WH. 1988 International collaborative study of hypertensive disorders of pregnancy. Geographic variation in the incidence of hypertension in pregnancy. Am J Obstet Gynecol 158:80–83. [PubMed: 2962500]

Osorio-Yanez C, Gelaye B, Miller RS, Enquobahrie DA, Baccarelli AA, Qiu C, et al. 2016 Associations of maternal urinary cadmium with trimester-specific blood pressure in pregnancy: Role of dietary intake of micronutrients. Biological Trace Element Research 174:71–81. [PubMed: 27129315]

Patel CJ, Bhattacharya J, Butte AJ. 2010 An environment-wide association study (ewas) on type 2 diabetes mellitus. PLoS One 5:e10746. [PubMed: 20505766]

Patel CJ, Rehkopf DH, Leppert JT, Bortz WM, Cullen MR, Chertow GM, et al. 2013 Systematic evaluation of environmental and behavioural factors associated with all-cause mortality in the united states national health and nutrition examination survey. International journal of epidemiology 42:1795–1810. [PubMed: 24345851]

Peterson R 2018 Bestnormalize: Normalizing transformation functions, r package version 1.2. 0.

Pinheiro T, Brunetto S, Ramos J, Bernardi J, Goldani M. 2016 Hypertensive disorders during pregnancy and health outcomes in the offspring: A systematic review. Journal of developmental origins of health and disease 7:391–407. [PubMed: 27168118]

Rhew IC, Vander Stoep A, Kearney A, Smith NL, Dunbar MD. 2011 Validation of the normalized difference vegetation index as a measure of neighborhood greenness. Annals of Epidemiology 21:946–952. [PubMed: 21982129]

Rolnik DL, Wright D, Poon LC, O'gorman N, Syngelaki A, de Paco Matallana C, et al. 2017 Aspirin versus placebo in pregnancies at high risk for preterm preeclampsia. New England Journal of Medicine 377:613–622. [PubMed: 28657417]

Rothman KJ, Greenland S, Lash TL. 2008 Modern epidemiology:Lippincott Williams & Wilkins.

Roux AVD, Merkin SS, Arnett D, Chambless L, Massing M, Nieto FJ, et al. 2001 Neighborhood of residence and incidence of coronary heart disease. New England Journal of Medicine 345:99–106. [PubMed: 11450679]

Rupasingha A, Goetz SJ, Freshwater D. 2006 The production of social capital in us counties. The journal of socio-economics 35:83–101.

Shoff C 2012 Disparities in hypertensive disorders of pregnancy across the levels and dimensions of rurality.

Sinisi SE, Van Der Laan MJ. 2004 Loss-based cross-validated deletion/substitution/addition algorithms in estimation.

Strand LB, Barnett AG, Tong S. 2011 The influence of season and ambient temperature on birth outcomes: A review of the epidemiological literature. Environ Res 111:451–462. [PubMed: 21333980]

Stuart JJ. Identifying women with a history of a hypertensive disorder of pregnancy: Values, challenges, and opportunities In: Proceedings of the Mayo Clinic Proceedings, 2018, Vol. 93Elsevier, 1695–1697.

Tooher J, Thornton C, Makris A, Ogle R, Korda A, Horvath J, et al. 2016 Hypertension in pregnancy and long-term cardiovascular mortality: A retrospective cohort study. American journal of obstetrics and gynecology 214:722. e721–722. e726. [PubMed: 26739795]

Tran TC, Boumendil A, Bussieres L, Lebreton E, Ropers J, Rozenberg P, et al. 2015 Are meteorological conditions within the first trimester of pregnancy associated with the risk of severe pre-eclampsia? Paediatric and perinatal epidemiology 29:261–270. [PubMed: 26053449]

U.S. EPA. 2014 Air quality data for the cdc national environmental public health tracking network. Available: http://www.epa.gov/heasd/research/cdc.html [accessed 21 March 2014].

University of Wisconsin School of Medicine and Public Health. 2015 Area deprivation index. Available: https://www.neighborhoodatlas.medicine.wisc.edu/.

USDA. 2014 Food access research atlas.

Valensise H, Vasapollo B, Gagliardi G, Novelli GPJH. 2008 Early and late preeclampsia: Two different maternal hemodynamic states in the latent phase of the disease. 52:873–880.

Van Buuren S 2018 Flexible imputation of missing data:Chapman and Hall/CRC.

van Middendorp D, Ten Asbroek A, Bio FY, Edusei A, Meijjer L, Newton S, et al. 2013 Rural and urban differences in blood pressure and pregnancy-induced hypertension among pregnant women in ghana. Globalization and health 9:59. [PubMed: 24228792]

Vinikoor-Imler LC, Gray SC, Edwards SE, Miranda ML. 2012 The effects of exposure to particulate matter and neighbourhood deprivation on gestational hypertension. Paediatric and perinatal epidemiology 26:91–100.

Wang IK, Tsai IJ, Chen PC, Liang CC, Chou CY, Chang CT, et al. 2012 Hypertensive disorders in pregnancy and subsequent diabetes mellitus: A retrospective cohort study. The American journal of medicine 125:251–257. [PubMed: 22340922]

White IR, Royston P, Wood AM. 2011 Multiple imputation using chained equations: Issues and guidance for practice. Statistics in medicine 30:377–399. [PubMed: 21225900]

Wikström A-K, Stephansson O, Cnattingius S. 2010 Tobacco use during pregnancy and preeclampsia risk: Effects of cigarette smoking and snuff. Hypertension 55:1254–1259. [PubMed: 20231527]

Wild CP. 2012 The exposome: From concept to utility. Int J Epidemiol 41:24–32. [PubMed: 22296988]

Wu P, Haththotuwa R, Kwok CS, Babu A, Kotronias RA, Rushton C, et al. 2017 Preeclampsia and future cardiovascular health: A systematic review and meta-analysis. Circulation: Cardiovascular Quality and Outcomes 10:e003497. [PubMed: 28228456]

Yackerson N, Piura B, Friger M. 2007 The influence of weather state on the incidence of preeclampsia and placental abruption in semi-arid areas. Clinical and experimental obstetrics & gynecology 34:27–30. [PubMed: 17447633]

Yoder SR, Thornburg LL, Bisognano JD. 2009 Hypertension in pregnancy and women of childbearing age. The American journal of medicine 122:890–895. [PubMed: 19786154]

Young C, Laurent O, Chung JH, Wu J. 2016 Geographic distribution of healthy resources and adverse pregnancy outcomes. Maternal and child health journal:1–7.

Zou H, Hastie T. 2005 Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology) 67:301–320.

Figure 1.

Figure 2.

Health Research Alliance Author Manuscript

Health Research Alliance Author Manuscript

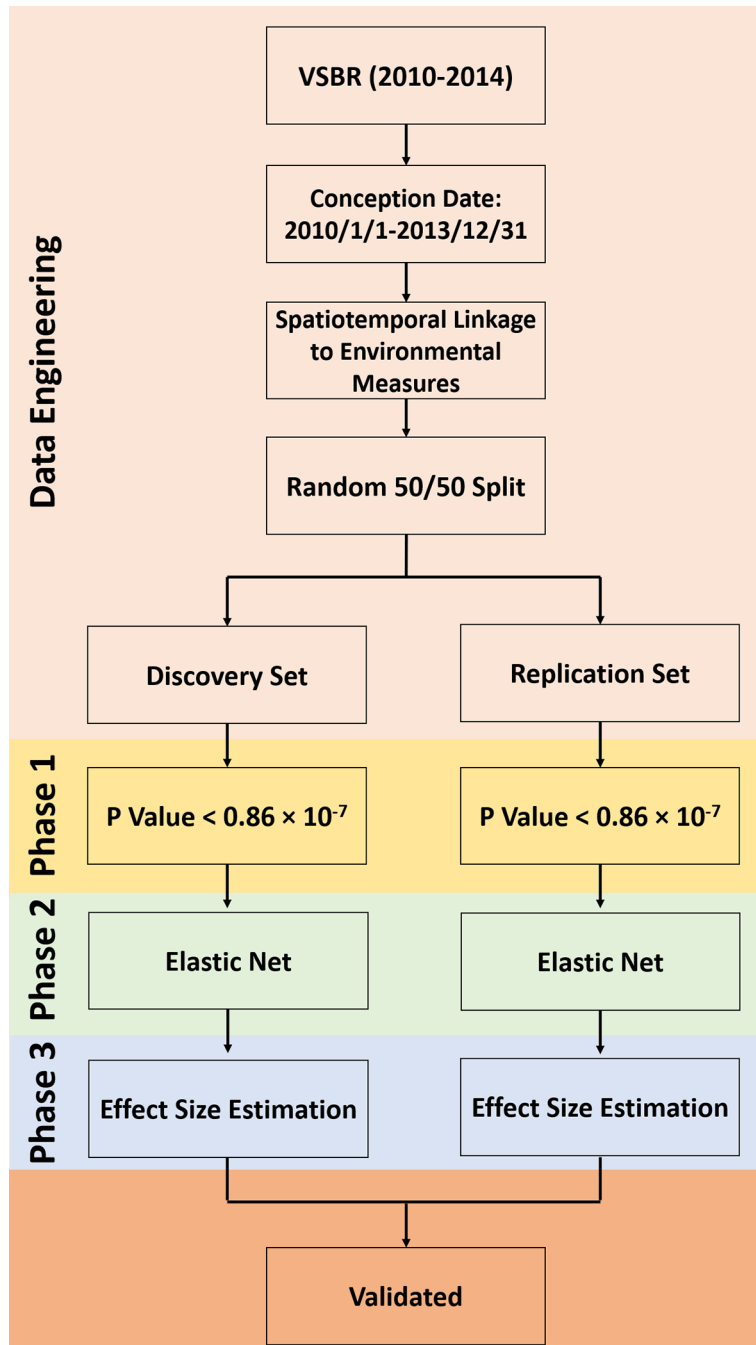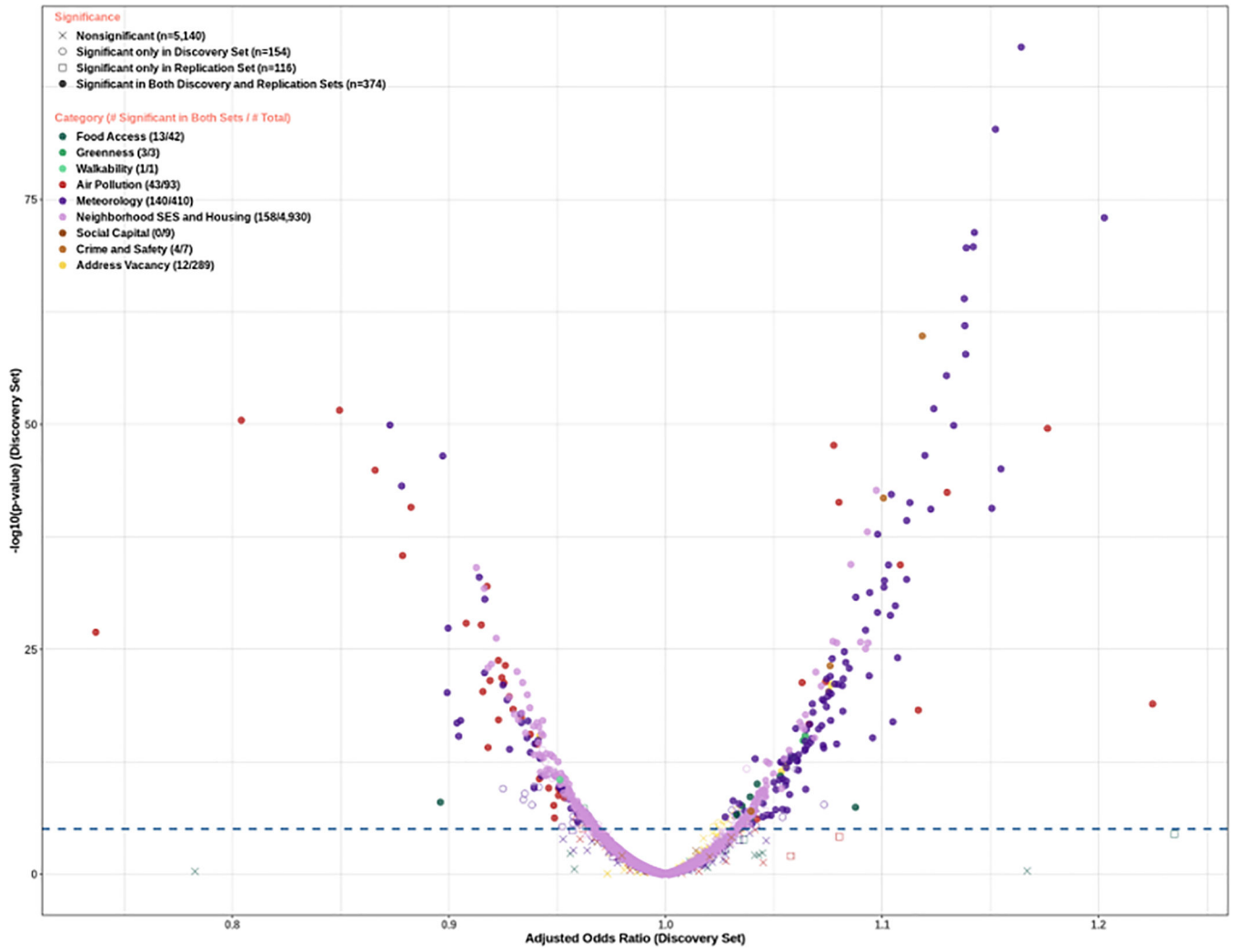|	1,1,1-Trichlorethane (T12)	2,2,4-Trimethypentane (T12)	Acrolein (T12)	Burglary rate (T12)	Forcible sex offenses rate (T12)	Maleic anhydride (T12)	Methyl tert-butyl ether (T12)	Omega at 100mb pressure level (T2)	Omega at 125mb pressure level (T12)	Proportion of individuals living in a place among workers 16 years and older (T12)	Proportion of individuals who speak English very well among those aged 18-64 years who speak Spanish (T12)	Propylene oxide (T12)
Propylene oxide (T12)	0.26	0.25	0.21	0.17	-0.07	-0.11	0.3	-0.01	-0.03	0.08	0.02	1
Proportion of individuals who speak English very well among those aged 18-64 years who speak Spanish (T12)	-0.06	-0.04	-0.01	0	0.04	-0.02	-0.02	0.04	0.03	0.01	1	0.02
Proportion of individuals living in a place among workers 16 years and older (T12)	0.19	0.36	0.08	0.05	0	-0.01	0.11	-0.08	-0.08	1	0.01	0.08
Omega at 125mb pressure level (T12)	-0.17	-0.21	0.04	-0.07	0.14	0	0.05	0.67	1	-0.08	0.03	-0.03
Omega at 100mb pressure level (T2)	-0.1	-0.17	0.06	0.01	0.15	0.01	0.02	1	0.67	-0.08	0.04	-0.01
Methyl tert-butyl ether (T12)	0.45	0.28	0.24	0.17	-0.07	-0.03	1	0.02	0.05	0.11	-0.02	0.3
Maleic anhydride (T12)	-0.08	-0.11	0.03	0.11	0.27	1	-0.03	0.01	0	-0.01	-0.02	-0.11
Forcible sex offenses rate (T12)	0	0.03	0.11	0.48	1	0.27	-0.07	0.15	0.14	0	0.04	-0.07
Burglary rate (T12)	0.35	0.29	0.19	1	0.48	0.11	0.17	0.01	-0.07	0.05	0	0.17
Acrolein (T12)	0.28	0.39	1	0.19	0.11	0.03	0.24	0.06	0.04	0.08	-0.01	0.21
2,2,4-Trimethypentane (T12)	0.42	1	0.39	0.29	0.03	-0.11	0.28	-0.17	-0.21	0.36	-0.04	0.25
1,1,1-Trichlorethane (T12)	1	0.42	0.28	0.35	0	-0.08	0.45	-0.1	-0.17	0.19	-0.06	0.26

value
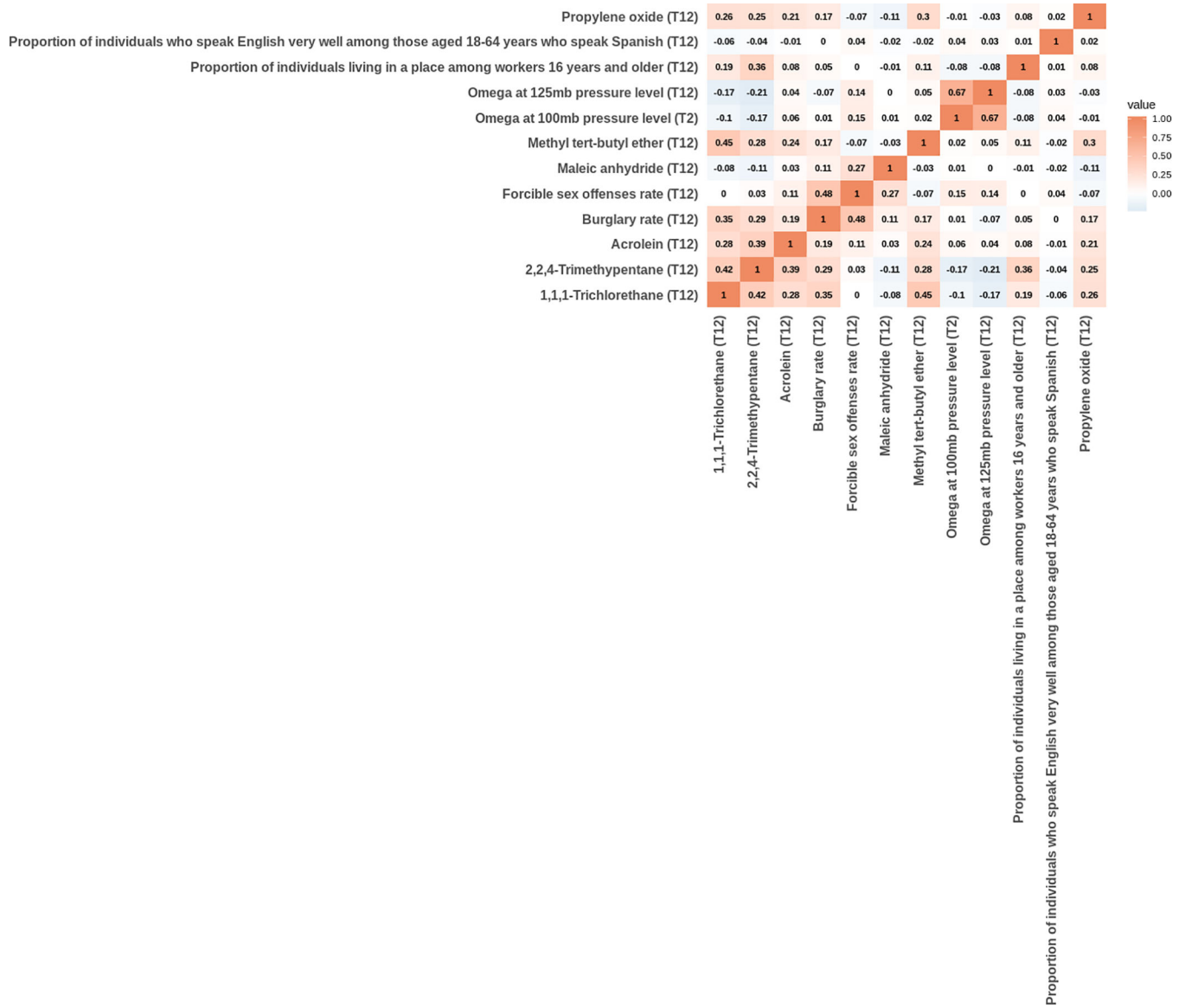1.00
0.75
0.50
0.25
0.00

Figure 3.

Table 1.

Summary of external exposome measures.

Category	Data source	Time	Spatial scale	Temporal scale	Spatial linkage	Temporal linkage[a]	Number of measures	Number of variables[b]
Natural environment								
Air toxicant	National Air Toxic Assessment, EPA	2005, 2011, 2014 (2010, 2012–2013 interpolated)	Census tract	1-year	250m buffer	T12	89	87
Fine particulate matter and ozone	Fused Air Quality Surface using Downscaling, EPA	2010–2014	Census tract	1-day	250m buffer	T1, T2, T12	2	6
Meteorology	North American Regional Reanalysis, National Centers for Environmental Prediction	2010–2014	Raster 32km	1-day	250m buffer	T1, T2, T12	257	410
Built environment								
Food access	Food Access Research Atlas, USDA	2010, 2015 (2011–2014 interpolated)	Census tract	1-year	250m buffer	T12	44	42
Greenness	Normalized Difference Vegetation Index, NASA's TERRA/MODIS	2010–2014	Raster 250m	16-day	250m buffer	T1, T2, T12	1	3
Walkability	WalkScore	2010–2014	Raster 0.0015 degree in lon/lat	Cross-sectional	250m buffer	T12	1	1
Social environment								
Address vacancy	Aggregated USPS Administrative Data on Address Vacancies, HUD	2010–2014	Census tract	3-month	250m buffer	T1, T2, T12	122	289
Crime and safety	Uniform Crime Reporting Program, FBI	2010–2014	County	1-year	250m buffer	T12	7	7
Social capital	Census Business Pattern	2010–2014	ZCTA5	1-year	250m buffer	T12	10	9
Sociodemographic status and housing	American Community Survey	2006–2015	Census block group	5-year	250m buffer	T12	4,930	4,930

[a]T1: trimester 1; T2: trimester 2; T12: trimesters 1 and 2.

[b]Number of variables after accounting for the different temporal linkages and removal of perfectly correlated variables (absolute value of correlation coefficient > 0.99).

Table 2.

Women's characteristics by HDP status among women with conception date during 2010–2013 in Florida, USA [n(%)].

Maternal characteristics	HDP 42,746 (5.2)	No HDP 776,653 (94.8)	Total 819,399 (100.0)
Age (years)			
<20	3,328 (5.6)	56,500 (94.4)	59,828 (7.3)
20–24	9,796 (5.0)	186,213 (95.0)	196,009 (23.9)
25–29	11,901 (5.1)	223,468 (94.9)	235,369 (28.7)
30–34	10,256 (5.0)	193,451 (95.0)	203,707 (24.9)
35–39	5,641 (5.7)	93,702 (94.3)	99,343 (12.1)
40	1,824 (7.3)	23,308 (92.7)	25,132 (3.1)
Missing	0 (0.0)	11 (100.0)	11 (0.0)
Race/ethnicity			
Non-Hispanic White	19,782 (5.3)	353,486 (94.7)	373,268 (45.6)
Non-Hispanic Black	10,248 (6.8)	140,833 (93.2)	151,081 (18.4)
Mexican	1,774 (3.7)	46,344 (96.3)	48,118 (5.9)
Puerto Rican	1,998 (4.4)	43,653 (95.6)	45,651 (5.6)
Cuban	2,267 (4.8)	44,624 (95.2)	46,891 (5.7)
Haitian	1,947 (6.2)	29,461 (93.8)	31,408 (3.8)
Others	4,730 (3.8)	118,251 (96.2)	122,981 (15.0)
Missing	0 (0.0)	1 (100.0)	1 (0.0)
Education			
<High school	5,665 (4.7)	115,720 (95.3)	121,385 (14.8)
High school or equivalent	13,372 (5.3)	240,268 (94.7)	253,640 (31.0)
Some college	9,210 (5.8)	148,858 (94.2)	158,068 (19.3)
College graduate	11,212 (5.2)	205,648 (94.8)	216,860 (26.5)
>College	3,049 (4.7)	61,845 (95.3)	64,894 (7.9)
Missing	238 (5.2)	4,314 (94.8)	4,552 (0.6)
Marital Status			
Not currently married	21,090 (5.4)	368,475 (94.6)	389,565 (47.5)
Currently married	21,653 (5.0)	408,137 (95.0)	429,790 (52.5)
Missing	3 (6.8)	41 (93.2)	44 (0.0)
WIC			
No	19,404 (5.1)	360,971 (94.9)	380,375 (46.4)
Yes	23,103 (5.3)	410,954 (94.7)	434,057 (53.0)
Missing	239 (4.8)	4,728 (95.2)	4,967 (0.6)
Smoking during pregnancy			
No	40,474 (5.3)	726,332 (94.7)	766,806 (93.6)
Yes	2,195 (4.3)	48,820 (95.7)	51,015 (6.2)
Missing	77 (4.9)	1,501 (95.1)	1,578 (0.2)
Pre-pregnancy BMI (kg/m^2)			
Underweight (<18.5)	932 (2.6)	35,599 (97.4)	36,531 (4.5)

Maternal characteristics	HDP 42,746 (5.2)	No HDP 776,653 (94.8)	Total 819,399 (100.0)
Normal (18.5–24.9)	12,450 (3.3)	360,687 (96.7)	373,137 (45.5)
Overweight (25.0–29.9)	10,982 (5.6)	185,610 (94.4)	196,592 (24.0)
Obese (30.0)	15,473 (9.3)	151,636 (90.7)	167,109 (20.4)
Missing	2,909 (6.3)	43,121 (93.7)	46,030 (5.6)
Parity			
Nulliparous	30,664 (5.1)	573,730 (94.9)	604,394 (73.8)
Parous	11,932 (5.6)	199,447 (94.4)	211,379 (25.8)
Missing	150 (4.1)	3,476 (95.9)	3,626 (0.4)
Season of conception			
Warm (June-November)	21070 (5.3)	379,545 (94.7)	400,615 (48.9)
Cool (December-May)	21,676 (5.2)	397,108 (94.8)	418,784 (51.1)
Year of conception			
2010	10,386 (5.1)	194,205 (94.9)	204,591 (25.0)
2011	10,648 (5.3)	192,038 (94.7)	202,686 (24.7)
2012	11,073 (5.4)	192,759 (94.6)	203,832 (24.9)
2013	10,639 (5.1)	197,651 (94.9)	208,290 (25.4)

Abbreviations: BMI, body mass index; HDP, hypertensive disorders of pregnancy; WIC, the Special Supplemental Nutrition Program for Women, Infants and Children.

Table 3.

Summary of the validated variables from the external ExWAS on HDP during 2010–2013 in Florida, USA (n=819,399).

| Exposure | | Transformation | Discovery set [b] | | | | | Replication set [b] | | | | |
| Variable [a] | Category | | Phase 1 | | Phase 2 | Phase 3 | | Phase 1 | | Phase 2 | Phase 3 | |
			OR (95% CI) [c]	p-value	Coefficient	OR (95% CI) [c]	p-value	OR (95% CI) [c]	p-value	Coefficient	OR (95% CI) [c]	p-value
2,2,4-trimethylpentane (T12)	Air toxicant	Box Cox (λ=0.24)	0.94 (0.92, 0.95)	$2.78×10^{-16}$	0.08	1.15 (1.08, 1.23)	$4.91×10^{-5}$	0.94 (0.93, 0.96)	$5.51×10^{-14}$	0.12	1.20 (1.13, 1.27)	$2.31×10^{-9}$
Omega at 125mb pressure level (T12)	Meteorology	Square root	1.06 (1.04,1.07)	$4.33×10^{-11}$	0.03	1.13 (1.04, 1.23)	$2.93×10^{-3}$	1.04 (1.03, 1.06)	$1.53×10^{-7}$	0.05	1.10 (1.01, 1.20)	$2.91×10^{-2}$
Acrolein (T12)	Air toxicant	$\log_{10}(x+0.001)$	1.07 (1.05, 1.08)	$2.11×10^{-17}$	0.08	1.09 (1.05, 1.13)	$1.25×10^{-6}$	1.07 (1.05, 1.08)	$1.51×10^{-18}$	0.06	1.06 (1.02, 1.10)	$2.09×10^{-3}$
Burglary rate (T12)	Crime and safety	Exponential	1.10 (1.09, 1.12)	$1.51×10^{-42}$	0.06	1.07 (1.05, 1.09)	$5.78×10^{-9}$	1.10 (1.08, 1.11)	$3.94×10^{-41}$	0.04	1.04 (1.02, 1.07)	$2.11×10^{-4}$
1,1,1-trichloroethane (T12)	Air toxicant	Box Cox (λ=0.12)	0.95 (0.93, 0.97)	$5.36×10^{-07}$	0.02	1.06 (1.01, 1.11)	$1.39×10^{-2}$	0.95 (0.93, 0.97)	$6.25×10^{-7}$	0.02	1.06 (1.01, 1.11)	$2.25×10^{-2}$
Proportion of individuals living in a place among workers 16 years and over (T12)	Sociodemographic status and housing	Square root	0.96 (0.95, 0.98)	$3.40×10^{-07}$	0.03	1.05 (1.01, 1.10)	$1.68×10^{-2}$	0.97 (0.95, 0.98)	$4.18×10^{-6}$	0.04	1.06 (1.03, 1.09)	$3.55×10^{-5}$
Maleic anhydride (T12)	Air toxicant	$\log_{10}(x+0.001)$	1.08 (1.07, 1.09)	$2.09×10^{-48}$	0.03	1.04 (1.01, 1.06)	$2.66×10^{-3}$	1.07 (1.06, 1.08)	$7.13×10^{-40}$	0.02	1.02 (1.01, 1.04)	$4.67×10^{-2}$
Forcible sex offenses rate (T12)	Crime and safety	Exponential	1.12 (1.10, 1.13)	$1.49×10^{-60}$	0.03	1.03 (1.01, 1.06)	$4.01×10^{-2}$	1.11 (1.10, 1.13)	$1.94×10^{-53}$	0.06	1.06 (1.03, 1.09)	$5.47×10^{-5}$
Propylene oxide (T12)	Air toxicant	$\log_{10}(x+0.001)$	0.95 (0.94, 0.97)	$3.09×10^{-09}$	0.02	1.02 (1.01, 1.04)	$3.11×10^{-2}$	0.96 (0.95, 0.98)	$5.87×10^{-6}$	0.02	1.03 (1.01, 1.05)	$1.18×10^{-2}$
Proportion of individuals who speak English very well among those aged 18–64 years who speak Spanish (T12)	Sociodemographic status and housing	No transformation	0.97 (0.95, 0.98)	$5.62×10^{-06}$	−0.02	0.98 (0.96, 0.99)	$7.19×10^{-3}$	0.97 (0.95, 0.98)	$9.57×10^{-7}$	−0.01	0.98 (0.96, 0.99)	$4.85×10^{-2}$

Exposure		Transformation	Discovery set[b]					Replication set[b]				
			Phase 1		Phase 2	Phase 3		Phase 1		Phase 2	Phase 3	
Variable[a]	Category		OR (95% CI)[c]	p-value	Coefficient	OR (95% CI)[c]	p-value	OR (95% CI)[c]	p-value	Coefficient	OR (95% CI)[c]	p-value
Omega at 100mb pressure level (T2)	Meteorology	$\log_{10}(x+0.01)$	1.05 (1.04, 1.07)	2.54×10^{-11}	−0.02	0.91 (0.87, 0.96)	2.07×10^{-4}	1.05 (1.04, 1.07)	1.78×10^{-11}	−0.02	0.95 (0.90, 0.99)	4.18×10^{-2}
Methyl tert-butyl ether (T12)	Air toxicant	$\log_{10}(x+0.001)$	0.80 (0.78, 0.83)	3.45×10^{-51}	−0.09	0.90 (0.85, 0.94)	1.44×10^{-5}	0.79 (0.77, 0.81)	6.56×10^{-59}	−0.05	0.94 (0.89, 0.99)	1.34×10^{-2}

[a]T1: trimester 1; T2: trimester 2; T12: trimesters 1 and 2.

[b]Models adjusted for age, race/ethnicity, education, marital status, WIC, smoking during pregnancy, pre-pregnancy BMI, parity, season of conception, and year of conception.

[c]Odds Ratio (OR) and 95% Confidence Interval (CI) for each standard deviation increase.