Identifying Urban Built Environment Factors in Pregnancy Care and Maternal Mental Health Outcomes

Yiye Zhang (yiz2014@med.cornell.edu)
Weill Cornell Medicine
Mohammad Tayarani
Cornell University
Shuojia Wang
Tencent (China)
Yifan Liu
Weill Cornell Medicine
Mohit Sharma
Weill Cornell Medicine
Rochelle Joly
Weill Cornell Medicine
Arindam RoyChoudhury
Weill Cornell Medicine
Alison Hermann
Weill Cornell Medicine
Oliver Gao
Cornell University
Jyotishman Pathak
Weill Cornell Medicine

Research Article

Keywords: Pregnancy care, postpartum depression, built environment

Posted Date: February 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-234334/v1

Abstract

Backgrounds: Environmental risk factors related to the built environment have been associated with women's mental health and preventive care. This study sought to identify built environment factors that are associated with variations in prenatal care and subsequent pregnancy-related outcomes in an urban setting.

Methods: In a retrospective observational study using machine learning, we characterized the types and frequency of events in prenatal care that are associated with the various built environment factors of the patients' residing neighborhoods. We hypothesize that, in comparison to women living in high-quality built environments, women who reside in low-quality built environments experience a different pattern of clinical events that may increase the risk for adverse outcomes. Using machine learning, we performed pattern detection to characterize the variability in prenatal care with respect to encounter types, clinical problems, and medication prescriptions. Structural equation modeling was used to test the associations among built environment, prenatal care variation, and pregnancy outcome. The main outcome is postpartum depression (PPD) diagnosis within 1 year following childbirth. The exposures were the quality of the built environment in the patients' residing neighborhoods. Electronic health records (EHR) data of pregnant women (n=8,949) who had live delivery at an urban academic medical center in 2015 to 2017 were included in the study.

Results: We discovered prenatal care patterns that were summarized into three common types. Women who experienced the prenatal care pattern with the highest rates of PPD were more likely to reside in neighborhoods with homogeneous land use, lower walkability, lower air pollutant concentration, and lower accessibility to retail stores after adjusting for age, neighborhood average education level, marital status, and income inequality.

Conclusions: In an urban setting, multi-purpose and walkable communities were found to be associated with a lower risk of PPD. Findings may inform urban design policies and provide awareness for care providers on the association of patients' residing neighborhoods and healthy pregnancy.

Background

The built environment, referring to the surroundings and physical artifacts of where humans live, is considered to be one of the five major social determinants of health (SDoH).(1) The built environment determines housing quality, mode of transportation, and exposure to pollutants, effectively influencing our way of life.(2) Poor built environment causes adverse effects on physical and mental health by disrupting sleep, hindering healthy life styles, and lowering access to healthcare.(3–5) There is a gender difference on the association between the built environment and health. Mullings et al. reported an increased risk of depression among female associated with living in an unplanned neighborhood characterized by inadequate sewer treatment, water supply, and dependable supply of electricity.(6) Furthermore, the Chicago Community Adult Health Study found the women's use of preventive care to be associated with objective and perceived neighborhood support and stressors such as odors, presence of trees, and noise levels.(7)

The existing literature motivated this study to examine the impact of the built environment on health and healthcare utilization among women, and particularly, the pregnant population.(8–10) Levels of prenatal care vary across the United States.(11–13) A substantial proportion of pregnant women, in particular those with a higher comorbidity burden or low health literacy, seek and depend on care provided by emergency departments (ED) rather than primary and obstetric care.(13–15) The lack of adequate prenatal care is considered to be a risk factor for poor pregnancy outcomes and lack of proper postpartum care for mothers and infants.(16) Previous studies have studied the built environment on maternal health and birth outcomes including birth weight, gestational age, Apgar score, and newborn intensive care unit admission rates.(5, 17) Yet, evidence is still accumulating on how the built environment affects the variability in prenatal care and maternal mental health outcomes. In particular, few studied the concurrent impacts of prenatal care and built environment on mental health outcomes. Existing studies have commonly relied on the subjective perceived measures obtained from interviews and questionnaires.(4, 7, 18) However, relying on subjective measurements may increase recall bias which occurs when some participants recall the exposure differently than others.

In this study, we hypothesize that the built environment, through influencing the accessibility to transportation, green space, safe neighborhood, and other urban structure, is associated with variability in prenatal care and subsequent maternal mental health outcomes.
Given findings from previous literature on the impact of the built environment on women’s mental health and use of healthcare, we defined postpartum depression (PPD) as our primary outcome.(19) PPD has been associated with increased infant mortality, higher rates of hospitalizations, impaired mother-child attachment, developmental problems in children, and increased stress within families.(20–23) The plethora of physical and psychological effects of PPD reported in previous studies include postpartum weight retention, reduced physical health, bodily pain, anxiety, low self-esteem, risky addictive behavior of substances, and suicide ideation.(24) The biological risk factors of PPD include genetic factors, age, pregnancy complications, medical illness, and smoking during pregnancy.(4, 25–27) The social, cultural, and environmental risk factors include income status, domestic violence, lack of social support, quantity and quality of green spaces, and residential noise pollution.(26, 28–32)

We tested our hypotheses by linking patients’ health data extracted from de-identified electronic health records (EHRs) with publicly available census-tract level data on the built environment. Routinely collected from clinical encounters, EHR data capture detailed longitudinal health data on health and health service utilizations. Increasingly, EHR data have been used as a source of longitudinal data in population health studies for its ability to provide detailed and rich health information within patient cohorts.(33) Leveraging a large cohort of nearly 9,000 women in New York City from 2015 to 2017, we applied machine learning algorithms to EHR data to identify patterns in prenatal care.(34) We then evaluated the relationships among prenatal care patterns, PPD incidence, and the built environment using structural equation modeling.(35) The association found may inform patients, care providers, and public health policy makers in supporting healthy pregnancy and new motherhood.

Methods

Study Setting

EHR Data

EHR data on 8,949 pregnant women from an urban academic medical center from 2015 to 2017 were extracted. The cohort inclusion and exclusion criteria are described in Fig. 1. We excluded patients whose ages were below 18 or above 45, had no encounter recorded in the EHR from 1 year prior to pregnancy to 1 year after delivery, or missing home locations information. We extracted patient information including gender, age, race, ethnicity, body mass index (BMI), marital status, outpatient and inpatient diagnoses, outpatient and inpatient prescription medication orders, and corresponding encounter dates from the EHR data. Patient age was calculated as the time difference between the birth date and first prenatal checkup date. The gestational week was calculated using the date of delivery and the specific gestational age at prenatal checkup. Marital status was defined as single (single, divorced, widowed, unknown), and married, as extracted from unstructured clinical notes using regular expression. The trimester of each event was determined using the difference in time between each event and delivery. All diagnoses were represented as Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) codes.(36) Anatomical Therapeutic Chemical (ATC) Classification System was used to standardize the specific drug prescription and dosage information.(37) The primary outcome of PPD was defined as having at least one diagnosis of depression within 1 year after childbirth based on SNOMED codes [see Additional file 1].

Built Environment Data

Accessibility to public transportation

Three indicators were defined to measure the accessibility to public transportation and active transportation facilities: the number of bus stops within 500-meter radius, the number of subway stations within the 500-meter radius, and the length of bike paths within the 500-meter radius. The spatial data on public transportation and bike facilities were obtained in shapefile formats from New York State.(38) We used ArcGIS 10.6 spatial analysis tools to count the number of bus stops and subway stations within each 500-meter radius around each patients' home location and also to measure the length of bike paths within the 500-meter radius.

Exposure to Traffic

We obtained traffic data from the New York activity-based travel demand model referred to as "New York Best Practice Model (NYBPM)."(39) The model predicts daily traffic volume in each roadway link for the different types of vehicles by two categories: light- (passenger vehicles and taxis) and heavy-duty (buses and trucks) vehicles for their different levels of health impacts.(40) The vehicle kilometer traveled (VKT) within the 500-meter radius was then calculated based on the distance that vehicle pollution concentration reaches the background level.(41) VKT is calculated by multiplying traffic volume by the distance of travel, representing the amount of traffic activity.

Land Use

Five indicators were defined to measure the role of land use: entropy-based land use mix (LUM) index, retail floor area ratio (RetFAR), street connectivity, and sidewalk availability. The variables measure the availability and variety of destinations within 500 meters of the subject's home location.

The land use data including information about land use class and parcel area at the parcel level were extracted from the parcel shapefile obtained from New York State.
(38) The LUM index within 500-m radius measures the heterogeneity of land use, such as residential, commercial, retail, and industrial, within the radius.(42) The LUM index ranges between 0 to 1, where 0 represents homogeneity and 1 represents maximum heterogeneity.(42) Higher LUM values indicate higher walkability of the area. The RetFAR is the retail building floor area divided by the retail land area within the 250-m radius.(42) Examples with higher and lower RetFAR are multi-floor departmental stores and open-style outlets, respectively. The number of intersections within the 500-meter radius is another land use indicator used to measure the walkability of the neighborhood.(43) The number of intersections was extracted from the transportation network developed for the NYBPM travel demand model. To calculate the sidewalk area within the 500-meter radius, we used the sidewalk shapefiles as a measure of the accessibility of subjects to the walking facilities.(39)

Air pollution

Average daily particulate matter ($PM_{2.5}$) and ozone (O_3) concentrations at the census tract level for the period of 2015–2017 were obtained from the Center for Air, Climate and Energy Solutions which applied Land Use Regression (LUR) models to estimate every subject's exposure to air pollution.(44) $PM_{2.5}$ and O_3 together could represent both regional background and hotspot air pollution levels.

Other Social Determinants of Health (SDoH)

Lastly, SDoH information at the census-tract (11-digit Federal Information Processing Standard code) level were extracted using the FACETS dataset.(45) Variables used in the analysis included census-tract level average percent of college degree, GINI index, felony rate, and uninsured percentage from American Community Survey, a binary indicator of low access to healthy food within half mile from the Food Access Research Atlas, United States Department of Agriculture, the population-weighted distance to closest 7 parks from the Centers for Disease Control and Prevention, and lastly walk score scales the from Rundle-Columbia Built Environment and Health Research Group.

Patterns of Prenatal Care

We extracted the health and healthcare utilization information during the prenatal period for each patient from the EHR data. Patients who had similar overall prenatal care patterns were categorized into clusters as having experienced generally similar prenatal events. The similarity between pairs of patients were measured using the longest common subsequence (LCS) distance. LCS measures the longest overlap that 2 sequences have in common; thus, larger LCS indicates a more similar course of the clinical events. In this study, we compared the sequence of each patient's clinical events (e.g., encounters, diagnoses, prescription medications) to others in the cohort to generate pairs of LCS distances. Based on the similarity, the categorization of patients was performed using the hierarchical clustering algorithm, a well-established machine learning method for detecting underlying clusters in a population.(34) The final number and size of the clusters were determined using Silhouette value.(34) This method was previously used to mine EHR data to identify health and healthcare utilization patterns among patients with chronic kidney disease, heart failure, and undifferentiated abdominal pain.(34, 46, 47) An example of the sequences used for categorization is given in the Additional file 2.

Because of the large number (n > 6,000) of unique clinical events recorded in the EHR data, we limited the pattern mining to focus on variables that were found to be most predictive of PPD in a related work preparatory to this study.(48) The list of variables, including complications during pregnancy and medication usage, are shown in Additional file 3. The cluster analysis was done in Python 3.6.5 and R 4.0.0.

Statistical Analysis

The distribution of study variables described in sections EHR Data and Built Environment Data (Table 1) were assessed within each identified cluster. Multivariate Imputation by Chained Equations (MICE) was used to address the missing value issue.(49) We further studied the relationship between prenatal care, as reflected by the cluster membership, the built environment characteristics, and incidence of PPD using structural equation models (SEMs).(35) Two SEMs were constructed for the primary and secondary outcomes separately. All independent variables were considered, but removed if there was multicollinearity as determined by variable inflation factor larger than 10. Statistical analysis was done using Stata/IC 16.0 and R 4.0.0. We applied Chi-square tests for categorical variables and analysis of variance (ANOVA) for continuous variables to compare the differences across clusters. P-value of 0.05 was used as the significance threshold.

Table 1
Descriptive statistics of the study cohort

Variables	Values
Demographics	
Age, mean (SD), year	33.69 (4.59)
Pre-pregnancy BMI, mean (SD), kg/m^2	23.77 (4.31)
Gestational Week, mean (SD), week	38.69 (2.09)
Race, No. (%)	
White	4409 (49.27)
Asian	1689 (18.87)
Black or African American	560 (6.26)
Other	976 (10.91)
Unknown	1315 (14.69)
Marital Status, No. (%)	
Single	1193 (13.33)
Married	7756 (86.67)
Cesarean Section, No. (%)	
Yes	1878 (20.99)
No	7071 (79.01)
Insurance, No. (%)	
Commercial	7519 (84.02)
Medicaid	1226 (13.70)
Other	204 (2.28)
Built Environment	
Number of bus stops within 500 m radius, mean (SD)	25.26 (10.0)
Number of subway stations within 500 m radius, mean (SD)	1.81 (1.83)
Parks Area within 500 m radius, mean (SD), m^2	463112.43 (660506.3)
Bike Path Length within 500 m radius, mean (SD), m	29070.94 (15172.89)
VKT of light vehicles within 500 m radius, mean (SD), 100,000 units	3283.87 (2242.98)

Variables	Values
VKT of heavy vehicles within 500 m radius, mean (SD), 10,000 units	3608.43 (2516.02)
LUM index within 500 m radius, mean (SD)	0.64 (0.17)
RetFar within 500 m radius, mean (SD)	0.24 (0.23)
Number of Intersections within 500 m radius, mean (SD)	12.06 (7.76)
Sidewalk Area within 500 m radius, mean (SD), 1000 m^2	907.77 (208.53)
Ozone Concentration, mean (SD), $\mu g/m^3$	46.56 (0.50)
$PM_{2.5}$ Concentration, mean (SD), $\mu g/m^3$	9.28 (0.47)
Percent of Colleges Degree, mean (SD), %	35.79 (11.49)
Average Poverty Rate, mean (SD), %	1.62 (2.15)
Average Respiratory Hazard Index, mean (SD)	4.51 (1.16)
Low Access to Healthy Food, No. (%)	297 (3.32)
Uninsured Percentage, mean (SD), %	8.26 (5.60)
Postpartum Depression	
Yes, No. (%)	273 (3.05)
Average number of ED visits per patient	
Pre-delivery (N = 3900, 43.58%), mean (SD)	0.74 (1.16)
Post-delivery (N = 482, 5.39%), mean (SD)	0.07 (0.31)

Results

Table 1 shows the descriptive statistics of the study cohort where continuous variables are presented as mean (standard deviation (SD)), and categorical variables are presented as N (% in total cohort). The average age of our patient population was 33.7 years (SD = 4.59). Nearly half (49.27%) of the patients were White, and majority were married (86.7%) and had Commercial insurances (84.1%). Over 3% of the cohort were diagnosed with PPD. A total of 3,922 (43.6%) and 482 (5.4%) patients had at least one ED visit pre- and post-delivery.

We identified 3 clusters with 1,955 (cluster 1), 4,188 (cluster 2), and 2,949 (cluster 3) patients, respectively, based on their clinical event sequences. For the primary outcome of PPD, 6.65% of the women in cluster 1 had a diagnosis of PPD within 1 year after childbirth, which was higher than clusters 2 (2.67%) and 3 (1.12%) ($P < .05$). Table 2 presents the distribution of demographics, medications, diagnoses, and built environment factors that were significantly different across the three clusters. The mean (SD) age across three clusters were 35.01 (4.73) years, 33.78 (4.29) years and 32.68 (4.66) years, respectively ($P < .001$). There were more unmarried patients in cluster 1 than the other two clusters ($P < .001$). In addition, the number of ED visits in both the pre- and post-delivery periods in the cluster 1 were significantly higher ($P < .05$) than the other clusters. In terms of medication prescriptions, we observed significantly higher rates of prescription medications in cluster 1, such as analgesics, antipyretics and opioids ($P < .001$). Further, more patients in cluster 1 had complications during pregnancy, unplanned pregnancies, high-risk pregnancy, abnormal glucose level, elderly primigravida and advanced maternal age gravidas than the other two clusters ($P < .001$). Additional file 2 showcases sequential patterns in the prenatal care identified from the study data.

Table 2
Associations between cluster membership and clinical variables used for clustering

Variables	Cluster			P-value
	1 (N = 1934)	2 (N = 4129)	3 (N = 2886)	
Demographics				
Age, mean (SD), year	35.01 (4.73)	33.78 (4.29)	32.68 (4.66)	< .001
Pre-pregnancy BMI, mean (SD), kg/m^2	24.24 (5.19)	23.55 (4.32)	23.77 (3.54)	< .001
Gestational Week, mean (SD), week	38.58 (2.12)	38.83 (1.92)	38.55 (2.26)	< .001
Race, no. (%)				
White	1078 (55.74)	2149 (52.05)	1182 (40.96)	< .001
Asian	280 (14.48)	679 (16.44)	730 (25.29)	
Black or African American	145 (7.50)	260 (6.30)	155 (5.37)	
Other	229 (11.84)	477 (11.55)	270 (9.36)	
Unknown	202 (10.44)	564 (13.66)	549 (19.02)	
Marital Status, no. (%)				
Single	348 (17.99)	578 (14.0)	267 (9.25)	< .001
Married	1586 (82.01)	3551 (86.0)	2619 (90.75)	
Average Poverty Rate, mean (SD), %	1.35 (1.83)	1.42 (1.87)	2.07 (2.61)	< .001
Cesarean Section, no. (%)				
Yes	510 (26.37)	833 (20.17)	535 (18.54)	< .001
No	1424 (73.63)	3296 (79.83)	2351 (81.46)	
Insurance, no. (%)				
Commercial	1603 (82.89)	3492 (84.57)	2424 (83.99)	.45

Variables	Cluster			P-value
	1 (N = 1934)	2 (N = 4129)	3 (N = 2886)	
Medicaid	283 (14.63)	552 (13.37)	391 (13.55)	
Other (Medicare, Self-pay, Unknown)	48 (2.48)	85 (2.06)	71 (2.46)	
ED Visits per patient				
Pre-delivery (within 1-year), mean (SD)	1.12 (1.54)	0.68 (1.01)	0.56 (0.97)	< .001
Post-delivery (within 6-months), mean (SD)	0.10 (0.37)	0.06 (0.29)	0.05 (0.28)	< .001
Medication Prescriptions				
Other Analgesics and Antipyretics, no. (%)	324 (16.75)	534 (12.93)	324 (11.23)	< .001
Opioids, no. (%)	285 (14.74)	323 (7.82)	243 (8.42)	< .001
Thyroid Preparations, no. (%)	291 (15.05)	273 (6.61)	84 (2.91)	< .001
Drugs for Functional Gastrointestinal Disorders, no. (%)	171 (8.84)	235 (5.69)	150 (5.2)	< .001
Antiemetics and Antinauseants, no. (%)	170 (8.79)	242 (5.86)	145 (5.02)	< .001
Other Plain Vitamin Preparations, no. (%)	172 (8.89)	252 (6.10)	83 (2.88)	< .001
Antihistamines for Systemic Use, no. (%)	185 (9.57)	234 (5.67)	83 (2.88)	< .001
Beta-lactam Antibacterials, Penicillins, no. (%)	175 (9.05)	245 (5.93)	81 (2.81)	< .001
Progestogens, no. (%)	284 (14.68)	156 (3.78)	42 (1.46)	< .001
Direct Acting Antivirals, no. (%)	143 (7.39)	187 (4.53)	70 (2.43)	< .001
Diagnoses				
Normal Delivery, no. (%)	1435 (74.2)	3346 (81.04)	2310 (80.04)	< .001
Primigravida, no. (%)	1206 (62.36)	2453 (59.41)	1024 (35.48)	< .001
Complication Occurring During Pregnancy, no. (%)	887 (45.86)	1439 (34.85)	605 (20.96)	< .001
Unplanned Pregnancy, no. (%)	641 (33.14)	1178 (28.53)	742 (25.71)	< .001

Variables	Cluster			P-value
	1 (N = 1934)	2 (N = 4129)	3 (N = 2886)	
Post-term Pregnancy, no. (%)	465 (24.04)	1116 (27.03)	532 (18.43)	< .001
Elderly Primigravida, no. (%)	674 (34.85)	935 (22.64)	360 (12.47)	< .001
High Risk Pregnancy, no. (%)	536 (27.71)	662 (16.03)	297 (10.29)	< .001
Abnormal Glucose Level, no. (%)	479 (24.77)	757 (18.33)	163 (5.65)	< .001
Advanced Maternal Age Gravida, no. (%)	416 (21.51)	675 (16.35)	222 (7.69)	< .001
Disorder of Pregnancy, no. (%)	342 (17.68)	499 (12.09)	276 (9.56)	< .001
Postpartum Depression				
Yes, no. (%)	130 (6.72)	110 (2.66)	33 (1.14)	< .001
No, no. (%)	1804 (93.28)	4019 (97.34)	2853 (98.86)	

Table 3 displays the results from the SEM for the outcome of PPD. Regarding the primary outcome, patients in clusters 1 (odds ratio = 6.3, P < .001) and 2 (odds ratio = 2.43, P < .001) are more likely to have a diagnosis PPD within 12 months after childbirth than women in cluster 3. Relative to cluster 3, patients in cluster 1 are more likely to have patients living in census tract that have lower PM 2.5 (odds ratio = 0.858, P = .02), lower retail floor area ratio (odds ratio = 0.882, P = .03), lower LUM (odds ratio = 0.508, P < .001), higher GINI (odds ratio = 4.317, P = 0.002), and higher college degree percentage (odds ratio = 4.401, P < .001). Patients are also more likely to be older in age (odds ratio = 1.115, P < .001) and not married (odds ratio = 0.404, P < .001). Relative to cluster 3, patients in cluster 2 are more likely to have patients living in census tract that have lower PM 2.5 (odds ratio = 0.890, P = 0.03), lower retail floor area ratio (odds ratio = 0.867, P = .001), lower GINI (odds ratio = 0.412, P = 0.02), and higher college degree percentage (odds ratio = 4.996, P < .001). Patients are also moderately more likely to be older in age (odds ratio = 1.046, P < .001) and not married (odds ratio = 0.560, P < .001). Race and insurance types (commercial, Medicaid, Other including Medicare) were not significantly associated with the cluster membership in the models although unadjusted association was significant.

Table 3
Built environment factors that are associated with cluster membership while controlling for social-demographic factors. OR: odds ratio

	Variable	OR	P-value
PPD	Cluster 1	6.3	< .001
	Cluster 2	2.43	< .001
Cluster 1 (vs. cluster 3)	Retail	0.882	.03
	PM2.5	0.858	.02
	Age	1.115	< .001
	Married	0.404	< .001
	LUM	0.508	< .001
	GINI	4.317	.002
	College	4.401	< .001
	_cons	0.069	< .001
Cluster 2 (vs. cluster 3)	Retail	0.867	.001
	PM2.5	0.890	.03
	Age	1.046	< .001
	Married	0.560	< .001
	LUM	0.749	.06
	GINI	0.412	.02
	College	4.996	< .001
	_cons	1.734	.33

Within each cluster, we further examined the characteristics of PPD cases as shown in Additional file 4. The association between PPD and the built environment factors were examined and shown in Additional file 5. The factors that were significantly associated with increased risk for PPD were the number of intersections within 500-m radius, the number of bus stops within 500-m radius, and retail floor area ratio, while adjusting for felony rates and GINI index which were also significant in the model.

Discussion

There were two major findings in this study. Three clusters of prenatal health and healthcare utilization patterns were discovered from a cohort of women whose pregnancies were managed entirely or partially in an urban academic medical center in 2015 to 2017. The distribution of the primary and secondary outcomes of PPD were significantly different across the clusters. Clinically, the clusters differed in maternal age, BMI, marital status, medication use, chronic conditions, and complications during pregnancy. In addition, we found that the cluster membership was associated with built environment factors related to walkability, access to retail resources, air quality, as well as neighborhood felony rates, and neighborhood income equality. These findings contribute to the growing body of evidence that the built environment in the community confers an impact on the trajectories of health and health service utilization during pregnancy.

The associations found between retail, land-use and the study outcomes among the pregnant cohort are novel and important contributions to the literature. Retail floor area ratio is indicative of pedestrian-orientated design and higher walkability. The mixed land use and more retail access may be a proxy for the connectedness of the neighborhood in providing community support to women. These community resources potentially lead to increased opportunities for social contact, lower stress levels, and higher physical activity levels, which is consistent with previous literature tying maternal mental health to green space.(9, 10) Air quality has been linked with adverse birth outcomes including preterm birth and miscarriages in previous literature.(9) However, we found that lower PM 2.5 concentration to be associated with clusters with higher PPD incidences in contrary to previous literature. In our urban study setting, PM 2.5 concentration is highest in the most affluent area and becomes lower as we move out to other parts of the study setting. Therefore, our findings on the association of poor air quality with higher incidence PPD case potentially reflect patient cohorts who are predominantly in or outside the most affluent part of the city who have better access to mental health reporting and care. Patterns learned from this study may inform expecting and new mothers, their care providers, as well as guideline and policy makers, to better prepare and navigate pregnancy and postpartum care. Additionally, our findings may have implications for policies during the current COVID-19 pandemic as our communities and their stores face significant changes.

There are limitations in the study. All diagnoses in the study were defined using diagnostic codes. Therefore, missed and under-diagnosis of health conditions during pregnancy, including PPD, is a crucial limitation. It is possible that this study missed PPD patients who did not disclose symptoms due to stigma against mental health, and patients who were diagnosed outside of our health system. The under- and mis-diagnosis may be more prevalent among women who live in low-income neighborhoods. Some of these limitations may be addressed in future work by patient interviews and questionnaires. Additionally, the application of natural language processing on unstructured clinical notes may allow us to elicit underdiagnosed and missed PPD as well as other conditions. Moreover, we were not able to address the possible reporting bias in our study population with respect to information such as race and marital status. Nearly 15% of the racial information was unknown from the EHR data. Future studies may explore the leveraging of patient-reported outcome data in overcoming this limitation. Furthermore, in analyzing the medication data, we did not consider the dose-response relationship between medications and the outcome as prescription fill information was not available. Detailed medication dose and frequency information can be analyzed in future work if pharmacy claims data become available. Lastly, while this study used data from a single health system in NYC, further work will aim to validate our findings using EHR data from other institutions and across different cities in the US.

Conclusion

We found that poor-quality built environment is associated with variability in prenatal care and maternal mental health outcomes in a large retrospective cohort study using EHR data.
Findings from this study may inform healthcare providers and public health policymakers in understanding modifiable risk factors that are associated with poor pregnancy care and outcomes.

Declarations

Ethics approval and consent to participate

This study was in accordance to guidelines of Weill Cornell Medicine, and was approved by the research ethics committee Weill Cornell Medicine Internal Review Board (protocol number: 1711018789). Meanwhile, the study was performed in accordance with the Declaration of Helsinki, including understanding the causes, development and effects of women's pregnancy-related disorder, improving preventive interventions, subject to ethical standards that promote and ensure respect for all human subjects and protect their health and rights, no possible harm to the environment, conducted only by individuals with the appropriate ethics and scientific education, training and qualifications. Also, the study includes information regarding funding, institutional affiliations, potential conflicts of interest and had no harm as consequence of participation in the research study.

Individual consent waiver for the study was obtained from Weill Cornell Medicine Internal Review Board (protocol number 1711018789).

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to its inclusion of patient health information protected by the Health Insurance Portability and Accountability Act but are available from the corresponding author on reasonable request.

Competing interests

YZ, AH, RJ, and JP have equity ownership at Iris OB Health, Inc.

MT, SW, MS, AR, YL, OG have no conflicts to disclose.

Funding

This study was funded by the Center for Transportation, Environment, and Community Health New Research Initiatives Fund and China Scholarship Council Fund.

Authors' contributions

YZ designed, analyzed, interpreted, and drafted the manuscript. MT, SW, and YL conducted the data analysis. MS conducted literature search. AH and RJ provided clinical interpretation of the results. AR provided statistical support. OG and JP provided guidance on study design.

Acknowledgements

Weill Cornell Medicine Information Technology Services, Research Informatics

References

1. Koh HK. A 2020 Vision for Healthy People. New England Journal of Medicine. 2010;362(18):1653–6.
2. Chaiyachati KH, Hom JK, Hubbard RA, Wong C, Grande D. Evaluating the association between the built environment and primary care access for new Medicaid enrollees in an urban environment using Walk and Transit Scores. Prev Med Rep. 2018;9:24–8.
3. Beutel ME, Brahler E, Ernst M, Klein E, Reiner I, Wiltink J, et al. Noise annoyance predicts symptoms of depression, anxiety and sleep disturbance 5 years later. Findings from the Gutenberg Health Study. Eur J Public Health. 2020;30(3):516–21.
4. Galea S, Ahern J, Rudenstine S, Wallace Z, Vlahov D. Urban built environment and depression: a multilevel analysis. J Epidemiol Community Health. 2005;59(10):822–7.
5. Emeruwa UN, Ona S, Shaman JL, Turitz A, Wright JD, Gyamfi-Bannerman C, et al. Associations Between Built Environment, Neighborhood Socioeconomic Status, and SARS-CoV-2 Infection Among Pregnant Women in New York City. JAMA: the journal of the American Medical Association. 2020.
6. Mullings JA, McCaw-Binns AM, Archer C, Wilks R. Gender differences in the effects of urban neighborhood on depressive symptoms in Jamaica. Rev Panam Salud Publ. 2013;34(6):385–92.
7. Veldhuis CB, Maki P, Molina K. Psychological and neighborhood factors associated with urban women's preventive care use. J Behav Med. 2020;43(3):346–64.
8. Guglielminotti J, Landau R, Wong CA, Li G. Patient-, Hospital-, and Neighborhood-Level Factors Associated with Severe Maternal Morbidity During Childbirth: A Cross-Sectional Study in New York State 2013–2014. Matern Child Health J. 2019;23(1):82–91.
9. McEachan RRC, Prady SL, Smith G, Fairley L, Cabieses B, Gidlow C, et al. The association between green space and depressive symptoms in pregnant women: moderating roles of socioeconomic status and physical activity. J Epidemiol Commun H. 2016;70(3):253–9.

10. Nichani V, Dirks K, Burns B, Bird A, Grant C. Green Space and Depression during Pregnancy: Results from the Growing Up in New Zealand Study. Int J Environ Res Public Health. 2017;14(9).
11. Glance LG, Dick AW, Glantz JC, Wissler RN, Qian F, Marroquin BM, et al. Rates Of Major Obstetrical Complications Vary Almost Fivefold Among US Hospitals. Health affairs. 2014;33(8):1330–6.
12. Grobman WA, Bailit JL, Rice MM, Wapner RJ, Varner MW, Thorp JM, et al. Can differences in obstetric outcomes be explained by differences in the care provided? The MFMU Network APEX study. American Journal of Obstetrics and Gynecology. 2014;211(2).
13. Farr SL, Dietz PM, Rizzo JH, Vesco KK, Callaghan WM, Bruce FC, et al. Health Care Utilisation in the First Year of Life Among Infants of Mothers With Perinatal Depression or Anxiety. Paediatr Perinat Ep. 2013;27(1):81–8.
14. Cunningham SD, Magriples U, Thomas JL, Kozhimannil KB, Herrera C, Barrette E, et al. Association Between Maternal Comorbidities and Emergency Department Use Among a National Sample of Commercially Insured Pregnant Women. Academic Emergency Medicine. 2017;24(8):940–7.
15. Kilfoyle KA, Vrees R, Raker CA, Matteson KA. Nonurgent and urgent emergency department use during pregnancy: an observational study. American journal of obstetrics and gynecology. 2017;216(2):181. e1-. e7.
16. D'Ascoli PT, Alexander GR, Petersen DJ, Kogan MD. Parental factors influencing patterns of prenatal care utilization. J Perinatol. 1997;17(4):283–7.
17. Akaraci S, Feng XQ, Suesse T, Jalaludin B, Astell-Burt T. A Systematic Review and Meta-Analysis of Associations between Green and Blue Spaces and Birth Outcomes. Int J Env Res Pub He. 2020;17(8).
18. Giurgescu C, Zenk SN, Templin TN, Engeland CG, Dancy BL, Park CG, et al. The Impact of Neighborhood Environment, Social Support, and Avoidance Coping on Depressive Symptoms of Pregnant African-American Women. Womens Health Issues. 2015;25(3):294–302.
19. Hahn-Holbrook J, Cornwell-Hinrichs T, Anaya I. Economic and Health Predictors of National Postpartum Depression Prevalence: A Systematic Review, Meta-analysis, and Meta-Regression of 291 Studies from 56 Countries. Front Psychiatry. 2017;8:248.
20. Jacques N, de Mola CL, Josephc G, Mesenburg MA, da Silveira MF. Prenatal and postnatal maternal depression and infant hospitalization and mortality in the first year of life: A systematic review and meta-analysis. J Affect Disorders. 2019;243:201–8.
21. Weobong B, ten Asbroek AHA, Soremekun S, Gram L, Amenga-Etego S, Danso S, et al. Association between probable postnatal depression and increased infant mortality and morbidity: findings from the DON population-based cohort study in rural Ghana. Bmj Open. 2015;5(8).
22. Field T. Postpartum depression effects on early interactions, parenting, and safety practices: A review. Infant Behav Dev. 2010;33(1):1–6.
23. Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. Lancet. 2014;384(9956):1800–19.
24. Moore Simas TA, Huang MY, Packnett ER, Zimmerman NM, Moynihan M, Eldar-Lissai A. Matched cohort study of healthcare resource utilization and costs in young children of mothers with

postpartum depression in the United States. Journal of medical economics. 2020;23(2):174–83.

25. Silverman ME, Reichenberg A, Savitz DA, Cnattingius S, Lichtenstein P, Hultman CM, et al. The risk factors for postpartum depression: A population-based study. Depress Anxiety. 2017;34(2):178–87.
26. Howard LM, Molyneaux E, Dennis CL, Rochat T, Stein A, Milgrom J. Non-psychotic mental disorders in the perinatal period. Lancet. 2014;384(9956):1775–88.
27. Chen HL, Cai JY, Zha ML, Shen WQ. Prenatal smoking and postpartum depression: a meta-analysis. J Psychosom Obstet Gynaecol. 2019;40(2):97–105.
28. OHara MW, Swain AM. Rates and risk of postpartum depression - A meta-analysis. Int Rev Psychiatr. 1996;8(1):37–54.
29. Zhang SM, Wang LS, Yang TB, Chen LZ, Qiu X, Wang TT, et al. Maternal violence experiences and risk of postpartum depression: A meta-analysis of cohort studies. Eur Psychiat. 2019;55:90–101.
30. Norhayati MN, Hazlina NHN, Asrenee AR, Emilin WMAW. Magnitude and risk factors for postpartum symptoms: A literature review. J Affect Disorders. 2015;175:34–52.
31. Feng XQ, Astell-Burt T. Residential green space quantity and quality and symptoms of psychological distress: a 15-year longitudinal study of 3897 women in postpartum. Bmc Psychiatry. 2018;18.
32. He SY, Smargiassi A, Low N, Bilodeau-Bertrand M, Ayoub A, Auger N. Residential noise exposure and the longitudinal risk of hospitalization for depression after pregnancy: Postpartum and beyond. Environ Res. 2019;170:26–32.
33. Schinasi LH, Auchincloss AH, Forrest CB, Roux AVD. Using electronic health record data for environmental and place based population health research: a systematic review. Ann Epidemiol. 2018;28(7):493–502.
34. Zhang Y, Padman R, Patel N. Paving the COWpath: Learning and visualizing clinical pathways from electronic health record data. Journal of biomedical informatics. 2015.
35. Huber C. Introduction to Structural Equation Modeling Using Stata. California Association for Institutional Research. 2014.
36. Odigie E, Lacson R, Raja A, Osterbur D, Ip I, Schneider L, et al. Fast Healthcare Interoperability Resources, Clinical Quality Language, and Systematized Nomenclature of Medicine-Clinical Terms in Representing Clinical Evidence Logic Statements for the Use of Imaging Procedures: Descriptive Study. JMIR Med Inform. 2019;7(2):e13590.
37. Ronning M, Blix HS, Harbo BT, Strom H. Different versions of the anatomical therapeutic chemical classification system and the defined daily dose - are drug utilisation data comparable? Eur J Clin Pharmacol. 2000;56(9–10):723–7.
38. State NY. The Official Website of New York State [Available from: https://www.ny.gov/.
39. Vovsha P, Petersen E, Donnelly R. Microsimulation in travel demand modeling: Lessons learned from the New York best practice model. Transportation Research Record: Journal of the Transportation Research Board. 2002(1805):68–77.

40. Pollution HEIPotHEoT-RA. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects: Health Effects Institute; 2010.
41. Karner AA, Eisinger DS, Niemeier DA. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. Environ Sci Technol. 2010;44(14):5334–44.
42. Frank LD, Sallis JF, Conway TL, Chapman JE, Saelens BE, Bachman W. Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. Journal of the American planning Association. 2006;72(1):75–87.
43. Sugiyama T, Leslie E, Giles-Corti B, Owen N. Associations of neighbourhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships? J Epidemiol Commun H. 2008;62(5).
44. Kim S-Y, Bechle M, Hankey S, Sheppard L, Szpiro A, Marshall J, editors. A Parsimonious Approach to National Prediction: Criteria Pollutants in the Contiguous US, 1979–2015. ISEE Conference Abstracts; 2018.
45. Cantor MN, Chandras R, Pulgarin C. FACETS: using open data to measure community social determinants of health. Journal of the American Medical Informatics Association. 2017;0(0).
46. Zhang Y, Padman R, Epner P, Bauer V, Solomonides A, Rao G. Identifying Diagnostic Paths for Undifferentiated Abdominal Pain from Electronic Health Record Data. AMIA Jt Summits Transl Sci Proc. 2018;2017:290-9.
47. Movahedi F, Kormos RL, Lohmueller L, Seese L, Kanwar M, Murali S, et al. Sequential pattern mining of longitudinal adverse events after Left Ventricular Assist Device implant. IEEE journal of biomedical and health informatics. 2019.
48. Wang S, Pathak J, Zhang Y. Using Electronic Health Records and Machine Learning to Predict Postpartum Depression. Studies in health technology and informatics. 2019;264:888–92.
49. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Meth Psych Res. 2011;20(1):40–9.

Figures

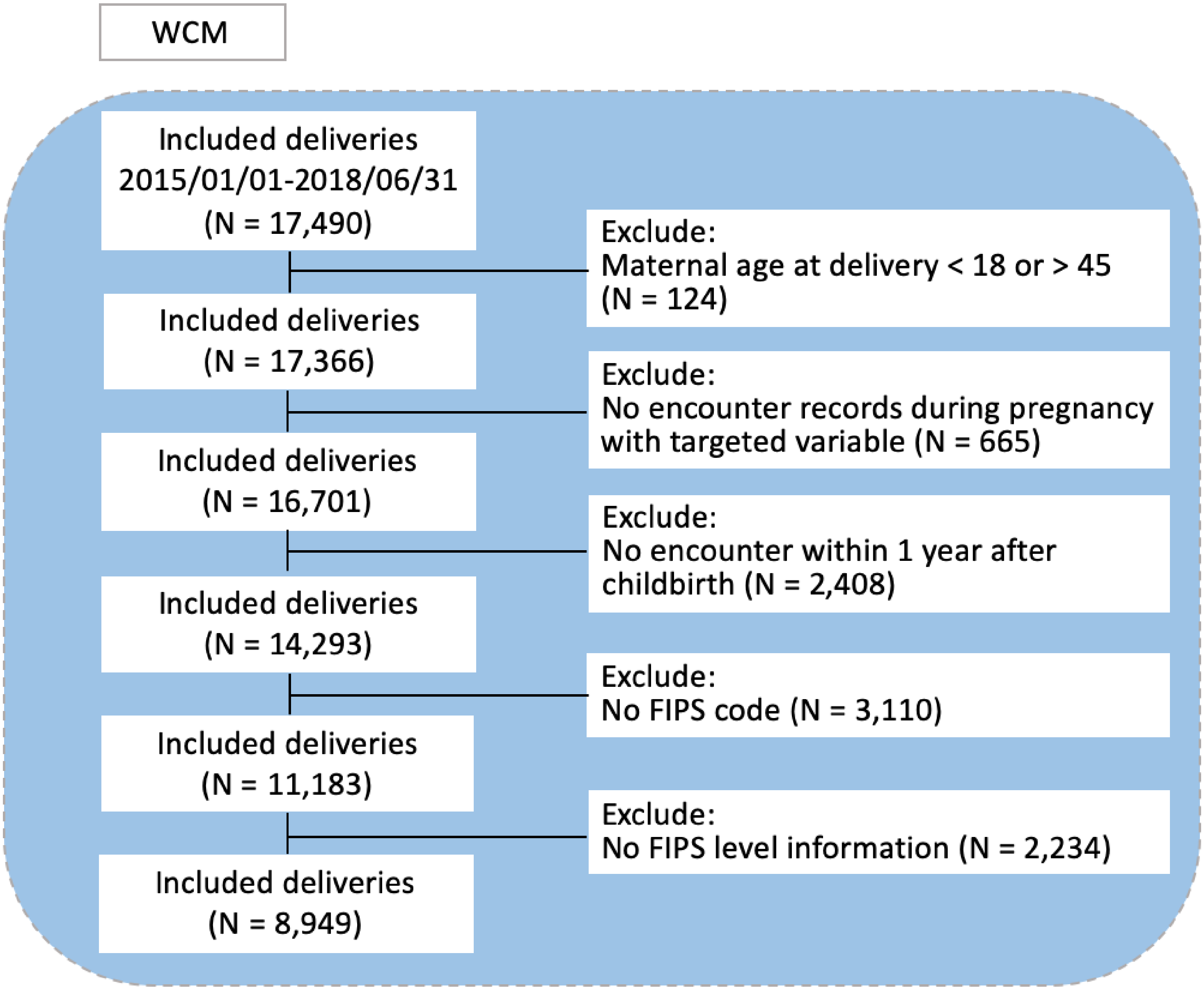

Figure 1

Study cohort inclusion and exclusion criteria

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile11.docx
- Additionalfile21.docx
- Additionalfile31.docx
- Additionalfile41.docx
- Additionalfile51.docx