2021

An Assessment of Kentucky Birth Records, Focusing on Early-Onset Hypertensive Disorders of Pregnancy, Environmental Metal Exposures, and Geocoding Precision, 2008-2017

Courtney J. Walker
University of Kentucky, court.j.walker@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0003-3274-7304
Digital Object Identifier: https://doi.org/10.13023/etd.2021.467

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Courtney J. Walker, Student

Dr. W. Jay Christian, Major Professor

Dr. Heather M. Bush, Director of Graduate Studies

AN ASSESSMENT OF KENTUCKY BIRTH RECORDS FOCUSING ON EARLY-ONSET HYPERTENSIVE DISORDERS OF PREGNANCY, ENVIRONMENTAL METAL EXPOSURES, AND GEOCODING PRECISION, 2008-2017

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Public Health
at the University of Kentucky

By

Courtney Janessa Walker

Lexington, Kentucky

Co-Directors: Dr. W. Jay Christian, Assistant Professor of Epidemiology
and Dr. Steven R. Browning, Associate Professor of Epidemiology

Lexington, Kentucky

https://orcid.org/0000-0003-3274-7304

ABSTRACT OF DISSERTATION

AN ASSESSMENT OF KENTUCKY BIRTH RECORDS FOCUSING ON EARLY-ONSET HYPERTENSIVE DISORDERS OF PREGNANCY, ENVIRONMENTAL METAL EXPOSURES, AND GEOCODING PRECISION, 2008-2017

Using live and stillbirth records from Kentucky (2008-2017), this dissertation assessed the county-level prevalence and geospatial patterns of early-onset hypertensive disorders of pregnancy (eHDP); examined the geocoding precision of addresses recorded on birth records, and evaluated the association between individual risk factors and environmental metal exposures on eHDP prevalence. After adjusting for maternal demographic factors and pre-existing health conditions, we observed that eHDP prevalence was 38% higher (aPR=1.38, 95%CI:1.16, 1.64) in counties with the highest prevalence of married women (>53.8%) compared to lower prevalence areas (<43%). We also found that counties with the highest prevalence of maternal obesity (>31.6%) had a 20% higher prevalence of eHDP(aPR=1.20, 95%CI:1.00, 1.44) compared to counties with lower obesity prevalence (<22.6%) after adjustment. We also identified two county-level clusters of eHDP in Appalachia. In the assessment of geocoding, we found that while address geocoding precision has improved over time, addresses of rural Black women were more likely to imprecisely geocode (aOR=1.41, 95%CI:1.22, 1.62) than rural White women. Adjusting for geocoding imprecision, we further assessed demographic and environmental factors associated with eHDP prevalence by augmenting records with census micro-block group toxicity concentration estimates of arsenic, cadmium, chromium, lead, and mercury from the Risk Screening Environmental Indicators (RSEI) model. Using a latent class analysis, we identified four classes of metal exposures. After adjusting for geocoding, imprecision, maternal demographics, and pre-existing health conditions, we found that women with a higher probability of lead and chromium exposure had a 22% higher prevalence of eHDP (aPR=1.22, 95%CI:1.04, 1.44) compared to women with a low probability of exposure to other considered metals. This study indicates an association between lead, chromium, and eHDP, even after adjusting for important covariates. Further research refining the use of RSEI scores and other exposures in association with eHDP is needed.

KEYWORDS: Early-onset hypertensive disorders of pregnancy, Risk Screening Environmental Indicators, Geocoding precision, Kentucky, Latent class analysis

Courtney Janessa Walker

11/16/2021

Date

AN ASSESSMENT OF KENTUCKY BIRTH RECORDS FOCUSING ON EARLY-ONSET HYPERTENSIVE DISORDERS OF PREGNANCY, ENVIRONMENTAL METAL EXPOSURES, AND GEOCODING PRECISION, 2008-2017

By

Courtney Janessa Walker

Dr. W. Jay Christian
Co-Director of Dissertation

Dr. Steven R. Browning
Co-Director of Dissertation

Dr. Heather Bush
Director of Graduate Studies

11/16//2021
Date

TABLE OF CONTENTS

LIST OF TABLES

Table 2.1 Total population count and percent of imprecisely geocoded Kentucky birth records among rural-urban classifications, 2008-2017 18
Table 2.2: Adjusted bivariate logistic regression of imprecise geocoding of Kentucky birth records by rural-urban classification 2007-2018 23
Table 3.1 Number of counties in each demographic group, and average eHDP prevalence by demographic group 34
Table 3.2 Unadjusted and adjusted prevalence of eHDP by demographic category 37
Table 3.3 Details of identified clusters of elevated eHDP prevalence, 2008-2017 39
Table 3.4 Prevalence of eHDP, marriage, maternal obesity, and maternal smoking per 1000 births in Kentucky, 2008-2017 42
Table 4.1 Indicators of fit for latent class analysis with two through four classes of emission exposures 57
Table 4.2 Demographic characteristics summary by latent metal class and the total population of primiparous mothers 2008-2017 59
Table 4.3 Bivariate and multivariable associations between demographic characteristics, environmental exposures, and class membership 62
Table A 1 Summary of Latent metal class and average posterior probability for class assignment 75

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Hypertensive disorders of pregnancy (HDP) accounted for 6.6% of maternal deaths from 2014-2017 in the US, making HDP a leading cause of morbidity and mortality in mothers and infants. [1,2] Subsets of this disorder include gestational hypertension (GH), pre-eclampsia (PE), pre-eclampsia with Hemolysis, Elevated Liver Proteins (PE + HELLP), superimposed pre-eclampsia, and eclampsia. These subsets are often progressive and include severe maternal outcomes such as placental abruption, pulmonary edema, stroke, and renal failure. [3,4] Long-term, women have an increased risk of cardiovascular and metabolic diseases. [5] Treatment options are limited, and often the only effective remedy is delivery of the placenta, which can be problematic in early-onset HDPs, as it increases the risk of poor outcomes for the infant. [5,6]

Between 1980 and 2003, the prevalence of HDP in the United States increased by 25%, with southern US states identified as having notably elevated prevalence. [7-9] In a one-year study assessing national trends, Kentucky was identified as having the 8th highest prevalence of HDP. [7] Overall, HDP is estimated to impact 8-10% of all pregnancies. Late-onset, or symptom manifestation after 34 weeks, is estimated to affect 2.7% of all pregnancies. Early-onset, characterized by symptoms manifesting before 34 weeks, is considered more severe and impacts approximately 0.38% of all pregnancies. [10,11]

Risk factors for HDP include primiparity, advanced maternal age, obesity, race, and use of infertility treatment; however, recent studies have suggested that risk factors for HDP, particularly PE, may need further refining, as there may be separate risk profiles for those who experience symptom earlier in their pregnancy. [6,12,13] Preliminary research suggests that environmental factors, such as exposure to arsenic, cadmium, chromium, lead, and mercury may also be associated with an increased risk of HDP. [14-21]

Long-term studies have found that arsenic (As), cadmium (Cd), chromium (Cr), lead (Pb), and mercury (Hg) are associated with cardiovascular disease (CVD) in non-pregnant adults; however, studies assessing acute effects, such as HDP, have had inconsistent findings. [22-24] Nonetheless, studies have demonstrated that close proximity to

industrial activity and dense road networks are associated with increased levels of toxic chemicals of concern [25-29] and increased risk of poor health outcomes such as cardiovascular disease and HDP. [30-33]

These conflicts, in part, may be due to the wide range of environmental sampling techniques and available data sources. [34] While some studies extrapolate exposure estimates based on sampling stations, and others may use Toxic Release Inventory data (TRI), a dataset of industrial emission estimates complied by the EPA, or conduct their own sampling. While the latter may be the gold standard, there may be limited resources to conduct robust environmental sampling. The TRI dataset, although robust, is an unadjusted dataset that contains estimated emissions, by volume, of specific chemicals tracked by the EPA and reported by mandated reporting facilities. To increase the usability of the TRI data, the EPA created the Risk Screening Environmental Indicators (RSEI) model. The RSEI model utilizes Toxic Release Inventory (TRI) emission reports to estimate the average toxicity concentrations of a given chemical in 810 m x 810 m grids across the United States. These grids are then overlayed with geographic boundary files to provide estimates at a census micro-block group (CMBG), census tract, ZIP code, and county level. Using the TRI emissions data for tracked chemicals from each emission site, RSEI creates toxicity concentration estimates adjusted for the fate and transport of the chemical throughout the environment, meteorological conditions, and site characteristics (when available) in the medium in which it is released. [35] Further details on RSEI methodology are available in the RSEI documentation. [35]

Emissions data with high spatial resolution may lead to the implicit assumption that the linked health records have a similar precision. However, the precision of geocoding coordinates can range substantially. Coordinates can be as precise as the rooftop of the address to as broad as the midpoint of a city. Further, previous studies evaluating geocoding precision have observed demographic characteristics associated with the record – such as race or ethnicity – or the population density of the residential area, factors often associated with disease status and severity are also associated with geocoding imprecision. [36,37] Exploratory models have found that imprecise geocoding can lead to misidentification of spatial clusters, biased results, and erroneous conclusions. [38,39]

The following chapters seek to address some of the described gaps by exploring the county-level prevalence and geographic distribution of eHDP, independent risk factors, and assessing the geocoding precision of maternal addresses obtained from birth records. Chapter two, " Geocoding precision of birth records from 2008 to 2017 in Kentucky, USA," focuses on the geocoding precision of birth records in Kentucky across rural-urban continuum codes (RUCCs) and sociodemographic factors associated with poor geocoding quality. In the third chapter, "County prevalence and geospatial trends of early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017," county-level estimates of eHDP are presented, clusters of eHDP in the state are discussed, and covariates associated with an increased prevalence of eHDP are identified. The fourth chapter, "A cross-sectional examination of the association of environmental metal exposure with early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017", explores the geospatial patterns of environmental emissions of arsenic, cadmium, chromium, lead, and mercury in Kentucky. Then, using a latent class analysis to identify subgroups of metal exposure, this study characterized the patterns within metal exposure classes and further assessed individual risk factors and environmental metal exposures on eHDP prevalence.

CHAPTER 2. GEOCODING PRECISION OF BIRTH RECORDS FROM 2008 TO 2017 IN KENTUCKY

2.1 Abstract

Maternal address information captured on birth records is increasingly being used to estimate residential environmental exposures during pregnancy. However, there has been limited assessment of the geocoding precision of birth records, particularly since the adoption of the 2003 standard birth certificate in 2015. To address this gap, this study evaluated the geocoding precision of live and stillbirth records of Kentucky residents over ten years – from 2008-2017. This study summarized the demographic characteristics of imprecisely geocoded records and, using a bivariate logistic regression, identified covariates associated with poor geocoding precision among three population density designations – urban, non-metro, and rural. We found that in urban areas, after adjusting for area deprivation, education, and the physical attributes of both parents, records for Black mothers had a 48% reduction in the odds of imprecise geocoding (aOR=0.52, 95% CI: 0.48, 0.56), while Black women in rural areas had a 96% increase in the odds of imprecise geocoding (aOR=1.96, 95% CI: 1.68, 2.28). This study also found that over the study period, rural and non-metro areas began with a high proportion of imprecisely geocoded records (38% in rural areas, 19% in non-metro), but both experienced an 8% decline in imprecisely geocoded records over the study period (aOR=0.92, 95% CI: 0.92, 0.94). This study shows that, while geocoding precision has improved in Kentucky, further work is needed to improve geocoding in rural areas and address racial and ethnic disparities.

2.2 Introduction

To assess retrospective environmental exposures during pregnancy, health researchers are increasingly employing US birth records as a source of residential information. However, studies evaluating birth record data quality predominantly focus on health information - the quality of address data are less frequently discussed in applied health research; although many of these studies employ geocoded address information to assess geospatial patterns of disease or evaluate residential environmental exposures [38-41].

Understanding the quality of residential address information, specifically the precision of geocoded addresses, is vitally important, as imprecise geocoding can result in records being assigned to an incorrect geographic location – leading to biased results and erroneous conclusions [38,39]. Of further concern is the evidence suggesting maternal demographic characteristics, notably race and ethnicity, are associated with the geocoding precision [37].

Geocoding assigns geographic coordinates to an address, using components such as address number, street name, city, state, or ZIP code [42-44]. Input addresses are compared to a reference dataset that contains street segments with address ranges and verified coordinates. Coordinates from the reference dataset are matched to the corresponding record of the input dataset [41]. These matches may be as precise as the '*rooftop*' of the actual residential structure or within a small range, *e.g.*, a location along a street segment based on the address range of that segment. Addresses missing components, such as the address or apartment number, or contain errors that prevent a sufficiently probable address match are assigned less precise coordinates that correspond to the mid-point (centroid) of the spatial resolution to which they were matched. Issues that can impede geocoding include spelling errors, use of special characters, neglected or incorrect suffixes (drive, lane, *etc*.), incorrect ZIP codes, or inaccurate or absent apartment numbers or complex names. Rural areas can pose additional challenges as rural or hired contractor routes have limited coordinate data. However, this may be less of an issue in future geocoding projects as the precision of e911 databases continue to expand and the increasingly common practice to assign specific addresses [36]. Overall, geocoding 85% or more addresses to either a rooftop or a street segment is considered a benchmark for high quality [45].

Many studies that employ geocoding to assign exposure status do not often report the proportion of records that geocoded imprecisely or the methods used to address imprecision and limit potential bias. [37,40,46] There has also been limited analyses of maternal characteristics that could be associated with poor geocoding precision – such as race, ethnicity, or education – which are factors that are also associated with disease status and severity. This cartographic confounding, or association between disease status and exposure, can bias spatial analyses, as imprecise records may be assigned incorrect

geographic locations which could result in systematic exposure misclassification if one group has a higher probability of imprecisely geocoding.

To address existing gaps in the literature, this study sought to describe the demographic characteristics of imprecisely geocoded records, identify covariates associated with imprecise geocoding, and explore regional variations in geocoding precision by examining maternal residential data from Kentucky birth records from 2008 through 2017. We believe that there may be a difference among metropolitan status regions, demographic characteristics and between the Appalachian and non-Appalachian region.

2.3 Materials and methods

2.3.1 *Data source and study population*

Kentucky Vital Statistics provided individual records of all live and stillbirths of Kentucky residents from January 1, 2008, through December 31, 2017. These records contained maternal addresses, maternal and paternal demographic information (age, race, ethnicity, education), self-reported marital status, maternal health information (height, pre-pregnancy weight), previous pregnancy history (parity), and characteristics of current pregnancy (birthplace, prenatal care, and plurality). Although live and stillbirth forms differ slightly, all variables used in this study were captured on both certificates. Non-singleton birth records were excluded to reduce the number of duplicate addresses (n=18,628), leaving 538,117 records for analysis.

2.3.2 Address-matched geocoding

We geocoded the maternal address using the ESRI Address Coder 10.7.1 (ESRI, Redlands, CA, USA). Address match score, or the level of agreement needed to match an address, was set at the program default of 93%. Geocoding outcomes were classified as precise (address point/street segment) or less precise (street name, ZIP code, city, or unknown). Addresses matched to a street name, city, or ZIP code were reviewed for spelling errors, inappropriate characters (for example, "0'Mally" rather than "O'Mally"), or neglected or incorrect designations (*e.g.*, road vs. boulevard) (N=30,879). After corrections and re-geocoding, 14.5% (N=4,608) of addresses that had initially been imprecisely geocoded were geocoded to a point or street segment.

2.3.3 *Covariates*

Maternal age in years was grouped into quartiles (<23, 23-25, 26-30, >30, unknown), as was paternal age (<24, 25-28, 29-33, >33, unknown). Due to small cell counts, race was classified into four categories (Black, White, Other, unknown). Maternal body mass index (BMI) was calculated using height and pre-pregnancy weight recorded on the birth record. Using standard classifications, BMI was then grouped into four categories: underweight (<18.5 kg/m^2), normal (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2) and obese (>30.0 kg/m^2). Appalachian designation, defined by the Appalachian Regional Commission (2020) (ARC), was assigned based on the geocoded county of residence. Rural-urban continuum codes (RUCC) from the United States Department of Agriculture were classified into three categories: rural (7-9), non-metro (4-6), and urban (1-3). (Alexander, 2011; United States Department of Agriculture, 2021) Other covariates considered in the model included ethnicity (Hispanic, non-Hispanic, unknown), education (less than high school, high school, some college and above, unknown), marital status (yes/no, or not stated), child birthplace (hospital, home birth, other), and parity (0, 1, 2+ previous births).

To classify prenatal care, we used the revised graduated prenatal care utilization index (R-GINDEX). Based on ACOG recommendations for the frequency of prenatal visits, this index assigns a score based on the percent of prenatal visits attended, adjusting for the date of the first prenatal visit (obtained from the birth record) and gestational age at delivery. [47] To adjust for local (i.e., micro block -group level) socioeconomic disadvantage, the area deprivation index (ADI) was used. [48,49] The ADI values, which are scaled from 1 to 10, were collapsed into deprivation quartiles: lowest (1-3), low (4-5), mid-range (ADI 6-7), and highest (8-10). Out of the 3,285 micro-block groups in Kentucky, 81 did not have an assigned ADI score due to low population counts or high group population quarters.

2.3.4 Analysis

Demographic characteristics were summarized with counts and row percentages by RUCC (rural, non-metro, and urban). A Chi-Square test was used to assess the distribution of frequencies of demographic covariates in geocoding precision. We used a multivariate binary logistic regression model for the final model and a backward

elimination method to identify covariates for the model. Our final model included maternal age, race, BMI, marital status, father's age, race, education, ethnicity, infant birthplace, parity, adequacy of gestational care, year of birth, ADI, and the Appalachian region. All non-spatial analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA). We created a line plot of the percentage of addresses imprecisely geocoded by rural-urban classification to visualize temporal trends.

2.4 Results

In this cross-sectional study, we found that 59% of birth records (n= 317,279) geocoded to an address point and 31% (n=166,752) geocoded to a street segment, yielding a precise geocoding rate of 90% (Figure 2.1).

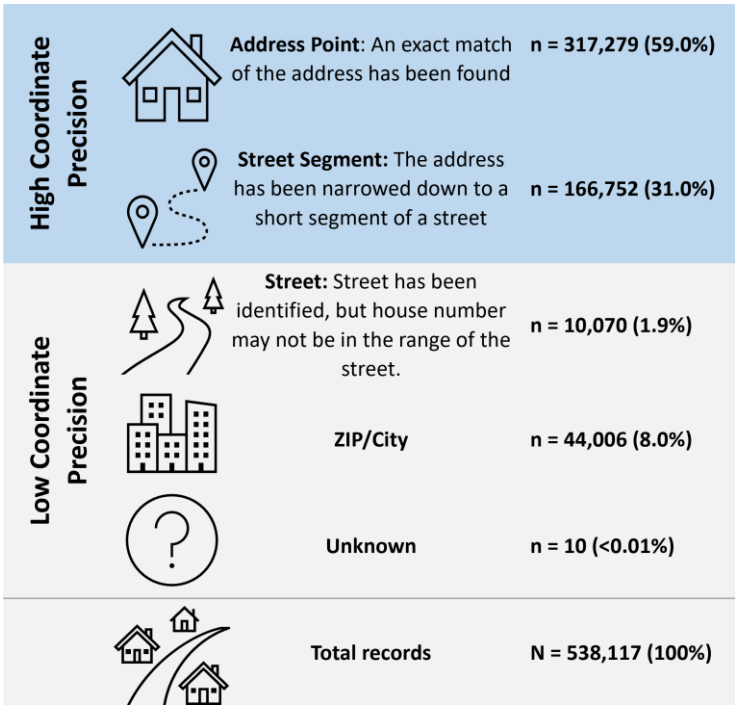


Figure 2.1 Description and summary of geocoding precision for Kentucky birth records, 2008-2017

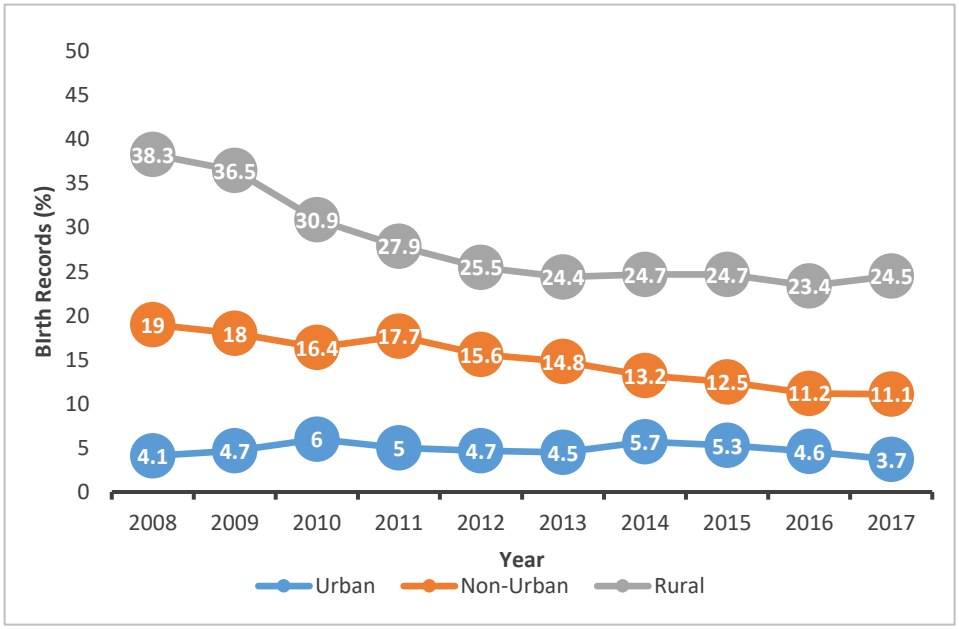


Figure 2.2 Prevalence of imprecise geocoding of Kentucky birth records by RUCC, 2008-2017

Figure 2.2 displays the proportion of records that geocoded imprecisely by RUCC over the study period. Rural areas experienced the sharpest decline (almost 10%) in imprecision. In contrast, urban areas had consistent geocoding imprecision over the study period. Non-metro areas experienced a slight decline.

Table 2.1 presents the number of records among each subgroup and the percent of records that geocoded imprecisely among all RUCCs and the entire sample. Overall, the population of Kentucky births occurred to women predominantly between the ages of 26-30 years old (n=154,789), who were non-Hispanic (n=509,425), White (n=450,425), had some college (n=293,878), and were married (n=311,817).

In all RUCCs, the prevalence of imprecise geocoding was highest among non-Hispanic women with less than a high school degree who were not married (Table 1). As educational attainment increased, the proportion of addresses that geocoded imprecisely decreased across all RUCCs. Among records with known maternal and paternal age, mothers < 23 years old and fathers who were < 24 years old had the highest proportion of imprecise records across all RUCCs as well. However, among RUCCs, the proportion of imprecisely geocoded records varied substantially among maternal and paternal racial groups. In urban areas, White mothers had a higher proportion of records not geocoded (5%); However, in non-metro and rural areas, a higher proportion of Black mothers did

not geocode precisely (non-metro: 16%, rural: 39%); which was similar to proportions observed within father's race. In urban areas, White fathers had the highest proportion of imprecise records (5%), while in non-metro and rural areas, we observed similar proportions as seen with Black mothers - 14% and 34%, respectively. Records with unknown characteristics, such as maternal age, paternal age, race, and ethnicity, did not have consistent geocoding precision patterns across RUCCs, except unknown paternal ethnicity, which had the largest proportion of imprecisely geocoded records across all RUCCs. In non-metro areas, mothers with unknown maternal and paternal age and maternal BMI had the largest prevalence of geocoding imprecision. In contrast, records with fathers of unknown age and education in rural areas had the largest proportion for their respective subgroups. Areas with low economic deprivation had higher geocoding imprecision in urban (6%) and non-metro (24%) areas; however, in rural areas, those with no assigned ADI had the largest proportion of imprecision (33%).

Table 2.1 Total population count and percent of imprecisely geocoded Kentucky birth records among rural-urban classifications, 2008-2017

	Urban N (%‡)	**Non-metro N (%‡)**	**Rural N (%‡)**	**Total N (%‡)**
Mother's Age (years)	*	*	*	*
<23	75540 (6.2)	54255 (16.7)	15590 (30.0)	145385 (12.7)
23-25	54212 (5.4)	32763 (15.0)	9194 (28.0)	96169 (10.8)
26-30	97265 (4.4)	45658 (13.8)	11866 (27.4)	154789 (8.9)
31-41	92288 (3.9)	34664 (14.1)	8712 (27.3)	135664 (8.0)
Unknown	2985 (5.1)	1381 (17.0)	473 (29.8)	4839 (10.9)
Mother's Race	*	*	*	*
White	250694 (5.3)	156255 (15.2)	43476 (28.1)	450425 (10.9)
Other	30619 (3.3)	6785 (11.4)	1283 (31.2)	38687 (5.7)
Black	40977 (3.3)	5681 (15.8)	1076 (39.4)	47734 (5.6)
Mother's Ethnicity	*	*	*	*
Hispanic	21743 (3.0)	4668 (11.1)	1010 (32.5)	27421 (5.4)
Not Hispanic	300547 (5.0)	164053 (15.1)	44825 (28.3)	509425 (10.3)
Mother's Education	*	*	*	*
Less than HS	47345 (6.6)	30443 (20.0)	10530 (31.5)	88318 (14.2)
High School	77265 (6.1)	54422 (15.5)	15463 (27.8)	147150 (11.9)
Some College +	193311 (3.9)	81418 (12.9)	19149 (27.2)	293878 (7.9)
Unknown	4369 (5.0)	2438 (14.2)	693 (27.8)	7500 (10.1)
Mother's BMI (kg/m ^{2})	*	*	*	*
Underweight (<18)	13293 (6.0)	8387 (18.1)	2459 (32.4)	24139 (12.9)
Normal (18-24.9)	137596 (4.5)	65378 (14.3)	17466 (27.6)	220440 (9.2)

Overweight (25-29.9)	79540 (4.8)	41048 (14.6)	11170 (28.4)	131758 (9.9)
Obese (>30)	79093 (5.4)	48884 (15.5)	13337 (28.5)	141314 (11.1)
Unknown	12768 (4.7)	5024 (18.1)	1403 (30.9)	19195 (10.1)
Marital Status	*	*	*	*
Yes	187167 (4.6)	97443 (14.1)	27207 (26.4)	311817 (9.5)
No or not stated	135123 (5.1)	71278 (16.3)	18628 (31.4)	225029 (10.8)
Father's Age	*	*	*	*
<24	48649 (6.4)	37215 (15.7)	10646 (28.4)	96510 (12.4)
25-28	57280 (5.0)	33017 (13.8)	8998 (26.1)	99295 (9.8)
29-33	78730 (3.9)	35148 (12.8)	9033 (25.4)	122911 (8.0)
>33	73810 (4.0)	31462 (14.5)	8420 (27.4)	113692 (8.7)
Unknown	63821 (5.7)	31879 (18.6)	8738 (34.9)	104438 (12.1)
Father's Race	*	*	*	*
White	206370 (5.1)	127882 (14.8)	36132 (27.2)	370384 (10.6)
Unknown	1738 (4.4)	1024 (14.4)	266 (20.3)	3028 (9.2)
Other	80693 (4.8)	33323 (16.1)	8292 (33.2)	122308 (9.8)
Black	33489 (3.3)	6492 (13.6)	1145 (34.1)	41126 (5.8)
Father's Ethnicity	*	*	*	*
Hispanic	19136 (3.1)	4839 (11.0)	1106 (31.2)	25081 (5.9)
Non-Hispanic	239191 (4.7)	132001 (14.5)	36138 (27.1)	407330 (9.9)
Unknown	63963 (5.8)	31881 (17.9)	8591 (33.7)	104435 (11.8)
Father's Education	*	*	*	*
Less than HS	32285 (6.7)	24573 (18.5)	9330 (28.3)	66188 (14.1)
High School	83242 (5.9)	60518 (14.9)	17306 (26.9)	161066 (11.5)
Some College +	140488 (3.4)	50757 (11.9)	10387 (26.9)	201632 (6.7)
Unknown	66275 (5.6)	32873 (17.6)	8812 (33.4)	107960 (11.5)
Birthplace	*	*	*	*
Hospital	318892 (4.8)	166600 (15.1)	44421 (29.1)	529913 (10.1)
Home	2947 (6.4)	1817 (9.4)	1241 (6.3)	6005 (7.3)
Other	451 (11.5)	304 (15.5)	173 (6.4)	928 (11.9)
R-GINDEX	*	*	*	*
Adequate (80%-109%,)	142625 (4.4)	80499 (13.6)	20110 (27.9)	243234 (9.4)
Inadequate (< 50%)	21281 (7.2)	13382 (18.8)	4549 (27.9)	39212 (13.6)
Intermediate (50-79%,)	114774 (4.8)	48528 (16.4)	14988 (28.7)	178290 (10.0)
Intensive (>110%)	23579 (5.6)	18811 (13.2)	4042 (26.9)	46432 (10.5)
Missing/No care	20031 (4.8)	7501 (19.2)	2146 (35.6)	29678 (10.6)
Parity	*	*		*
0	130713 (4.6)	68068 (15.1)	17662 (28.9)	216443 (9.9)
1	102585 (4.7)	54893 (14.5)	14497 (28.4)	171975 (9.9)
2 +	88992 (5.3)	45760 (15.6)	13676 (27.7)	148428 (10.5)

ADI	*	*	*	*
Lowest (1-3)	143561 (3.2)	25357 (4.6)	1189 (20.9)	170107 (3.6)
Mid-low (4-5)	65378 (5.0)	36980 (6.2)	4735 (20.1)	107093 (6.1)
Mid-high (6-7)	59047 (6.4)	41800 (14.1)	11865 (20.0)	112712 (10.7)
Highest (8-10)	51115 (5.9)	61757 (24.4)	27098 (33.3)	139970 (19.4)
Missing ADI	3189 (27.8)	2827 (33.7)	948 (45.6)	6964 (32.6)
Stillbirths			*	
No or not stated	320552 (4.8)	167697 (15.0)	45569 (28.5)	533818 (10.1)
Yes	1738 (4.4)	1024 (14.4)	266 (20.3)	3028 (9.2)
Appalachian	*	*	*	*
Not Appalachia	307406 (4.7)	79709 (6.6)	12658 (15.2)	399773 (5.4)
Appalachia	14884 (8.2)	89012 (22.6)	33177 (33.5)	137073 (23.6)
Total	322290 (4.8)	168721 (15.0)	45835 (28.4)	536846 (10.1)

BMI: Body Mass Index R-GINDEX: Revised Graduated Prenatal Care Index, ADI: Area Deprivation Index ‡ Percent imprecisely geocoded

Figures 2.3 illustrates the county-level geocoding imprecision across the state. For additional context, Figure 4 displays the RUCC for each county and identifies larger cities (>50,000 inhabitants) and the state capital, Frankfort. The Appalachian region had multiple counties characterized by many imprecise geocodes. Counties with a high population density and their contiguous neighbors, on the other hand, had much higher percentages of addresses that geocoded precisely.

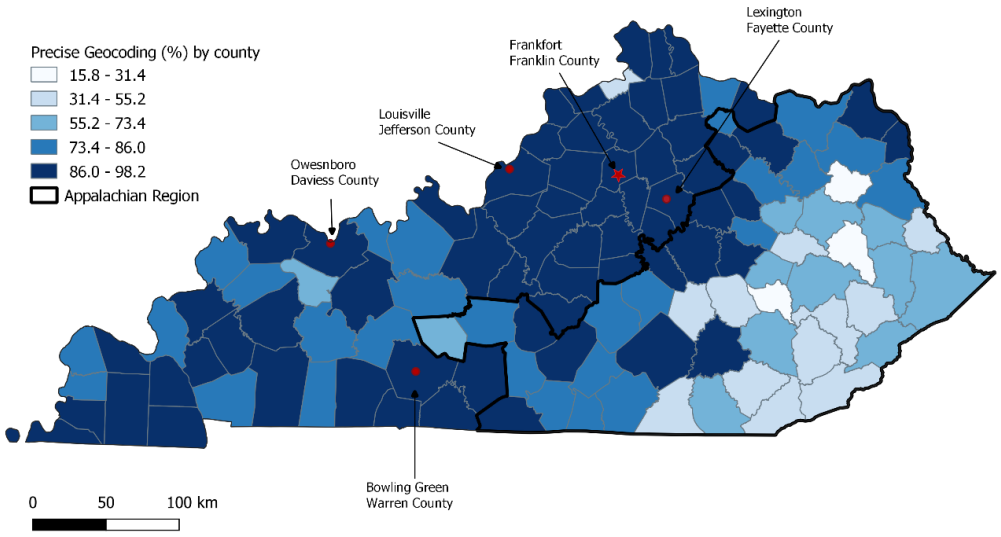


Figure 2.3 Percent of Kentucky birth record addresses that geocoded to an address point or street segment by county, 2008-2017

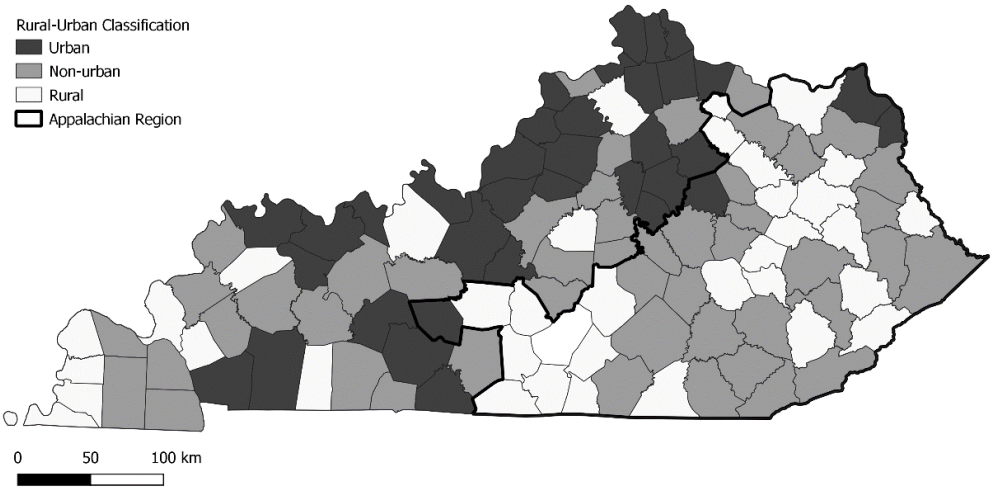


Figure 2.4 Kentucky rural-urban classification areas (RUCC), 2010

The multivariable logistic regression, presented in Table 2, displays notable differences in the odds of imprecise geocoding among RUCCs. In urban areas, Black mothers had a 48% reduction in the odds of imprecise geocoding, adjusting for other covariates (aOR=0.48, 95% CI: 0.48, 0.56); however, in rural areas, Black women had 96% higher odds of imprecise geocoding (aOR=1.96, 95% CI: 1.68, 2.28). Maternal ethnicity also varied by RUCC, with non-Hispanic mothers in urban areas having 74% higher odds of imprecise geocoding (aOR=1.74, 95% CI: 1.58, 1.94), 20% higher in non-

metro areas (aOR=1.20, 95% CI: 1.06, 1.38), but 14% lower in rural areas, although the latter outcome was not statistically significant (aOR=0.86, 95% CI: 0.70, 1.04). Patterns within father's race and ethnicity were slightly different, with Black fathers in urban areas having a lower odds of imprecision (aOR=0.72, 95% CI: 0.66, 0.78); however, non-metro areas had a higher, although non-significant odds of imprecise geocoding, compared to White fathers (aOR=1.04, 95% CI: 0.94, 1.12). In rural areas, like maternal race, the odds of imprecision were higher with Black fathers than White (aOR=1.32, 95% CI: 1.14, 1.54).

As maternal age increased, the odds of imprecision decreased in urban areas, with those between the ages of 31-41 years experiencing the lowest odds (aOR=0.84, 95% CI: 0.78, 0.90). However, in non-metro and rural areas, as maternal age increased, the odds of imprecision increased, with those aged 31-41 years reporting 10% higher odds of imprecision in non-metro areas (aOR=1.10, 95% CI: 1.04, 1.16) and 12% in rural areas, although non-significant (aOR=1.12, 95% CI: 0.90, 1.42). Paternal age also significantly impacted geocoding precision. In non-metro areas, paternal age >33 years had a significant 6% increase in the odds of imprecise geocoding (aOR=1.06, 95% CI: 1.00, 1.12). In non-metro (aOR=1.24, 95% CI: 1.12, 1.34) and rural (aOR=1.26, 95% CI: 1.10, 1.42) areas, unknown father's age was associated with a significantly increased odd of imprecise geocoding, compared to the fathers <24 years. Compared to mothers who were married, non-married mothers had 30% lower odds of imprecise geocoding (aOR=0.70, 95% CI: 0.68, 0.74), although, in rural areas, those who were married had 10% higher odds (aOR=1.10, 95% CI: 1.06, 1.16). Inadequate care was associated with an increased odds of geocoding imprecision in all RUCCs, compared to adequate care started in the first trimester. Although non-significant in rural areas. In urban areas, intensive care had substantially higher odds of imprecision (aOR 1.26, 95% CI: 1.18, 1.34).

Across all RUCCS, increased maternal education was associated with a decreased odds of imprecision compared to those with no high school degree. Father's education was protective against imprecision in urban and non-metro areas.

Compared to mothers who were married, mothers were not married had 30% lower odds of imprecise geocoding (aOR=0.70, 95% CI: 0.68, 0.74), although, in rural

areas, those who were married had 10% higher odds (aOR=1.10, 95% CI: 1.06, 1.16). Inadequate care was associated with an increased odds of geocoding imprecision in all RUCCs, compared to adequate care started in the first trimester. Although non-significant in rural areas. In urban areas, intensive care had substantially higher odds of imprecision (aOR 1.26, 95% CI: 1.18, 1.34).

The impact of socioeconomic disadvantage, as measured by ADI, is also noteworthy concerning geocoding precision. Those who geocoded in areas without ADI scores had the highest odds of not being properly geocoded in urban (aOR=12.22, 95% CI: 11.22, 13.32) and non-metro areas (aOR=8.5, 95% CI: 7.68, 9.40). Over the study period, the odds of imprecise geocoding remained unchanged in urban areas. However, in non-metro areas, the odds declined by 8% (OR=0.92, 95% CI: 0.92, 0.94) and by 8% in rural areas (OR: 0.92, 95% CI: 0.92, 0.94).

Table 2.2: Adjusted bivariate logistic regression of imprecise geocoding of Kentucky birth records by rural-urban classification 2007-2018

	Urban	**Non-metro**	**Rural**
	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)
Mother's Age (years)			
<23	Reference	Reference	Reference
23-25	0.94 (0.88, 0.98)	1.02 (0.98, 1.06)	1.06 (0.98, 1.12)
26-30	0.88 (0.82, 0.92)	1.04 (1.00, 1.10)	1.06 (1.00, 1.14)
31-41	0.84 (0.78, 0.90)	1.10 (1.04, 1.16)	1.12 (1.02, 1.22)
Unknown	0.98 (0.82, 1.18)	1.16 (0.98, 1.36)	1.12 (0.9, 1.42)
Mother's Race			
White	Reference	Reference	Reference
Other	0.80 (0.74, 0.86)	1.16 (1.04, 1.3)	1.16 (0.98, 1.38)
Black	0.52 (0.48, 0.56)	1.42 (1.3, 1.54)	1.96 (1.68, 2.28)
Mother's Ethnicity			
Hispanic	Reference	Reference	Reference
Not Hispanic	1.74 (1.58, 1.94)	1.20 (1.06, 1.38)	0.86 (0.70, 1.04)
Mother's Education			
Less than HS	Reference	Reference	Reference
High School	0.9 (0.86, 0.94)	0.84 (0.80, 0.86)	0.82 (0.78, 0.88)
Some College +	0.76 (0.72, 0.8)	0.84 (0.80, 0.88)	0.84 (0.80, 0.90)
Unknown	0.84 (0.72, 1)	0.68 (0.60, 0.78)	0.68 (0.54, 0.82)
Mother's BMI (kg/m^2)			
Normal (18-24.9)	Reference	Reference	Reference
Underweight (<18)	1.14 (1.06, 1.24)	1.14 (1.08, 1.22)	1.12 (1.02, 1.24)

Overweight (25-29.9)	1.08 (1.04, 1.12)	1.06 (1.02, 1.08)	1.04 (1.00, 1.10)
Obese (>30)	1.14 (1.10, 1.18)	1.1 (1.06, 1.14)	1.06 (1.00, 1.12)
Unknown	1.02 (0.94, 1.12)	1.28 (1.18, 1.38)	1.22 (1.08, 1.38)
Marital Status			
Yes	Reference	Reference	Reference
No or not stated	0.70 (0.68, 0.74)	1.02 (1.00, 1.06)	1.10 (1.06, 1.16)
Father's Age			
<24	Reference	Reference	Reference
25-28	0.90 (0.86, 0.96)	0.96 (0.92, 1.00)	0.92 (0.86, 0.98)
29-33	0.80 (0.76, 0.86)	0.96 (0.90, 1.00)	0.88 (0.82, 0.94)
>33	0.88 (0.82, 0.94)	1.06 (1.00, 1.12)	0.96 (0.90, 1.06)
Unknown	0.84 (0.76, 0.94)	1.24 (1.12, 1.34)	1.26 (1.10, 1.42)
Father's Race			
White	Reference	Reference	Reference
Other	0.70 (0.64, 0.74)	0.78 (0.72, 0.84)	0.94 (0.84, 1.06)
Black	0.72 (0.66, 0.78)	1.04 (0.94, 1.12)	1.32 (1.14, 1.54)
Unknown	0.50 (0.38, 0.64)	0.68 (0.54, 0.82)	0.52 (0.38, 0.72)
Father's Ethnicity			
Hispanic			
Non-Hispanic	1.28 (1.16, 1.44)	0.98 (0.86, 1.12)	0.82 (0.68, 0.98)
Unknown	2.38 (2.02, 2.78)	1.28 (1.10, 1.50)	0.88 (0.70, 1.08)
Father's Education			
Less than HS	Reference	Reference	Reference
High School	0.84 (0.78, 0.88)	0.90 (0.86, 0.94)	0.96 (0.90, 1.04)
Some College +	0.54 (0.50, 0.56)	0.84 (0.80, 0.88)	1.06 (0.98, 1.14)
Unknown	0.72 (0.64, 0.82)	0.78 (0.70, 0.86)	0.92 (0.78, 1.08)
Birthplace			
Hospital	Reference	Reference	Reference
Home	0.88 (0.76, 1.04)	0.66 (0.56, 0.78)	0.18 (0.14, 0.24)
Other	2.28 (1.70, 3.06)	0.78 (0.56, 1.08)	0.16 (0.08, 0.30)
Parity			
0			
1	1.06 (1.02, 1.12)	0.96 (0.94, 1.00)	1.02 (0.96, 1.06)
2+	1.14 (1.10, 1.20)	1.00 (0.96, 1.04)	1.02 (0.96, 1.08)
R-GINDEX			
Adequate (80%-109%)			
Inadequate (< 50%)	1.40 (1.32, 1.48)	1.12 (1.06, 1.18)	1.04 (0.96, 1.12)
Intermediate (50-79%)	1.04 (1.00, 1.08)	1.10 (1.06, 1.14)	1.00 (0.96, 1.06)
Intensive (>110%)	1.26 (1.18, 1.34)	0.98 (0.94, 1.04)	1.00 (0.92, 1.08)
Missing/No care	0.98 (0.92, 1.06)	1.18 (1.10, 1.26)	1.36 (1.24, 1.52)
ADI			
Lowest (1-3)			
Mid-high (6-7)	1.38 (1.32, 1.46)	1.30 (1.20, 1.40)	0.58 (0.48, 0.68)

Highest (8-10)	1.78 (1.70, 1.88)	2.66 (2.50, 2.84)	0.42 (0.36, 0.50)
Mid-low (4-5)	1.68 (1.60, 1.76)	4.46 (4.20, 4.76)	0.68 (0.58, 0.80)
Missing ADI	12.22 (11.22, 13.32)	8.50 (7.68, 9.40)	1.20 (0.98, 1.46)
Appalachian Region			
Non-Appalachian			
Appalachia	1.38 (1.30, 1.48)	3.16 (3.04, 3.26)	2.72 (2.56, 2.90)
Year	1.00 (1.00, 1.02)	0.92 (0.92, 0.94)	0.92 (0.92, 0.94)

HS= High School; Body Mass Index; ADI=Area Deprivation Index R-GINDEX: Revised Graduated Prenatal Care Index

2.5 Discussion

To more adequately characterize potential limitation for studies that employ geocoding to assess disease prevalence and exposure status, further review is needed. This study used Kentucky birth records from 2008-2017 to characterize the prevalence of geocoding imprecision among RUCCs, identify characteristics associated with imprecise geocoding, and assess the geospatial distribution of geocoding imprecision in the state. We found that the proportion of addresses that geocoded imprecisely in rural and non-metro areas declined over the study period. We also found maternal and paternal race, age, ethnicity, and education, along with marital status, prenatal care, and ADI were significantly associated with imprecision, but the magnitude and direction varied among RUCCs. Further, we identified that rural regions, particularly in the Appalachian region, had the highest proportion of imprecisely geocoded addresses.

Racial and ethnic disparities have been found in other studies assessing geocoding precision in birth records. In a study using birth data from Florida, Ha and colleagues reported that Black women had a higher odds of not geocoding than White women [37]. Gilboa and colleagues reported a higher proportion of Latinas in non-geocoded records in an assessment of the precision of Texas birth records and birth defects registry data [50]. In this study, although almost all records were geocoded, we found both parents' race and ethnicity influenced the odds of precise geocoding after adjusting for other factors, such as urban density, that impact the odds of precise geocoded. Although records of Black women and men were more likely to geocode precisely in urban areas, this may be due to the allocation of GIS resources in communities that have a higher prevalence of pollution and contamination – such as historically redline non-white communities [34]. In rural areas, which are less likely to have historically redlined areas as they are not as racially

diverse, Black women had a lower geocoding precision. Further study exploring explanations, and solutions, in rural areas is needed.

Rural residency has been cited as an important factor in the inability to geocode in Florida and Virginia [37,51,52]. However, certain conventions in rural areas have shifted, *e.g.*, introducing emergency 911 telephone (E911) systems, allowing for more accurate identification of addresses[53]. This study used the national address coder provided through the software manufacturer, but future work exploring other sources, including historical records, for geospatial coordinates is needed.

This study also found that educational and economic factors were also associated with geocoding imprecision. Both maternal and paternal education was associated with a decreased odds of imprecision, compared to the respective parent not completing high school. The protective effect of education, which often serves as a proxy for income, may reflect financial means to choose more attentive hospitals or reflect form literacy; however, further work is needed. We also found that addresses geocoded in areas that were not assigned an ADI score, which was due to having either a low population count or high group population quarters, had a much higher odds of imprecise geocoding than areas with low economic deprivation. Across all RUCCs, however, an increase in economic deprivation was associated with an increased odds of imprecise geocoding. Given that economic deprivation is often associated with poor health outcomes, it is clear that further work to improve GIS capabilities in these areas is needed.

Cartographic confounding, or the association demographic factors (such as race and rurality), with both the outcome of interest and geocoding precision, is an under-discussed but pressing concern in studies of health geography. [51] In an assessment of research employing geographic information systems (GIS) methodologies Cromley [40] noted that 25% of articles assessing health literature did not describe their techniques. in a meta-analysis of electronic health records using GIS methods, Shinasi and colleagues noted that less than one-third of journal articles described their geocoding process. This lack of transparency in disclosing geocoding methodologies leaves other health researchers unable to evaluate a study's merits and limitations fully.

Limitations of this study include noted variations among hospitals in vital statistics collection and reporting. We were also unable to assess the records for accuracy or identify addresses changes during the pregnancy or preconception period, as only current residential information is captured on the birth record. As we used the currently available geocoding reference file through ESRI, older records may contain addresses that no longer exist due to redevelopment, and would not have been able to be precisely identified. However, a noted strength of this study was the inclusion of all singleton birth records – still and live birth – in Kentucky, over ten years. Further, we were able to assess maternal and paternal demographic factors, economic factors and regional differences among all Kentucky births.

Future work may consider using state specific geocoders, and employ historical records to improve geocoding of older records, and explore solutions for improving geocoding for all races and ethnicities, particularly in rural regions.

CHAPTER 3. COUNTY PREVALENCE AND GEOSPATIAL TRENDS OF EARLY-ONSET HYPERTENSIVE DISORDERS OF PREGNANCY IN KENTUCKY, 2008-2017

3.1 Abstract

Early-onset hypertensive disorders of pregnancy (eHDP) are associated with more severe maternal and infant outcomes than the later-onset disease; however, eHDP has had a limited evaluation of prevalence and geospatial trends. In this study, we used Kentucky certificates of live and stillbirth to assess county-level spatio-temporal trends and covariates associated with an increased prevalence of eHDP. We found that after adjusting for race (Black %), educational attainment (% completed college), maternal smoking (%) that counties with the highest obesity prevalence (≥31.6%) had a 20% increase in eHDP prevalence compared to counties with the lowest obesity prevalence (<22.6%) (aOR=1.20, 95% CI: 1.00, 1.44). We also found counties with the highest proportion of primiparous mothers ≥34 years old (>6.1%) had a 26% increase in the prevalence of eHDP (OR=1.26, 95% CI: 1.04, 1.50), compared to counties with the lowest prevalence (<2.5%). We further identified two county-level clusters of elevated rates of eHDP in the Appalachian region. These trends may reflect poor reproductive literacy and poor community health.

3.2 Introduction

Hypertensive disorders of pregnancy (HDP), a progressive disease of pregnancy which includes diagnoses such as gestational hypertension (GH), preeclampsia (PE), and eclampsia and accounted for 6.6% of maternal deaths from 2014-2017 and impacts approximately 8-10% of US pregnancies each year. [1,2] Maternal complications of HDP can include pulmonary edema, renal failure, stroke, and death. [4,54,55] Women who have early-onset disease onset not only have an increased risk of experiencing an HDP in future pregnancies, they have a higher risk of cardiovascular disease, with a younger age of disease onset compared to women who had late-onset PE or normotensive pregnancies. [56] Pharmaceutical and behavioral prevention and interventions are incredibly limited. For women perceived as "high risk" a daily low dose aspirin can be recommended to reduce risk, but after the onset of symptoms delivery is often the only

option, however, premature delivery increases the infant's risk of poor health outcomes. [5] Black women, primiparity, extreme maternal age, or those with pre-existing health conditions such as diabetes, chronic hypertension, or are obese, are at the highest risk of developing HDP. [57-59] However, recent literature suggests that subsets of HDP, such as PE, may encompass multiple conditions with different etiologies and risk factors, that in part, may be distinguishable by gestational age at disease onset. [59,60]

Between 1980 and 2003, the estimated prevalence of HDP in the United States increased by 25%, and several studies have identified the Southern US states as having elevated HDP prevalence. [7-9] In a more recent one-year study assessing national trends, Kentucky was identified as having the 8th highest prevalence of HDP. [7] Regional differences observed in geospatial trends of eHDP have been attributed to geographic variation in health behaviors and the prevalence of preexisting conditions. [61,62] However, our understanding of geospatial trends and population-level covariates associated with HDP and eHDP prevalence is limited.

Although our understanding of HDP has improved, further exploration, particularly with early-onset disease, is needed. Using Kentucky birth records from 2008-2017, this longitudinal ecological study will describe the distribution of eHDP among covariate subgroups, identify county-level covariates associated with increased prevalence eHDP, and assess spatiotemporal trends of eHDP.

3.3 Methods

We obtained IRB approval from both the Kentucky Cabinet for Health and Family Services and the University of Kentucky. We followed Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines in reporting study results. [63]

3.3.1 Study Population and Outcome Ascertainment

The Kentucky Department of Vital Statistics provided individual records for all live and still births to self-identified Kentucky residents from January 1, 2008, through December 31, 2017. These records contained addresses, maternal information (age, marital status, race, education, ethnicity, number of previous births, height, and pre-

pregnancy weight), and pregnancy characteristics (gestation length, cigarettes smoked before and during each trimester of pregnancy, prenatal care, parity, number of previous pregnancies, and complications of pregnancy). Although live and stillbirth forms differ slightly, all variables used in this study were captured on both certificates. [64] We used individual records to identify singleton births to primiparous women (ages 13-50 years) between 20 to 45 weeks gestation and geocoded in Kentucky. Records that did not geocode (n=3) or indicate the mother had pre-existing chronic hypertension (n = 3,854) were excluded, as HDP and chronic hypertension are mutually exclusive on birth records. [64,65] Individual-level data were assessed for patterns in missingness using PROC MI in SAS.

The birth form provides separate checkboxes for chronic hypertension, gestational hypertension, and eclampsia. Gestational hypertension includes transient hypertension, PE, and HELLP (Hemolysis, Elevated Liver enzyme levels, and Low Platelet syndrome). [65] Early-onset HDP (eHDP) was defined as check positive for HDP on the birth certificate and gave birth between 20 through 34 weeks of gestation.

We used Rural-Urban Continuum Codes (RUCC) obtained from the United States Department of Agriculture to assess population density. [66] Cartographic boundary files were obtained from the United States Census Bureau. [67] Appalachian status, defined by the Appalachian Regional Commission (ARC), was based on the geocoded maternal county of residence. [68]

3.3.2 Data cleaning and dataset preparation

Maternal age was calculated by subtracting the maternal date of birth from the child's date of birth. Women $\geq$34 years or older were classified as being of advanced maternal age. Maternal obesity status was derived using the mother's pre-pregnancy weight and height. A body mass index (BMI) of 30 kg/m^2 or greater was classified as obese. [69] Smoking status was derived from the number of cigarettes reported being consumed during each trimester of pregnancy. Women who reported cigarette use during every trimester until delivery were classified as smokers. Women who reported no cigarette consumption or did not report cigarette consumption after the first trimester were considered non-smokers, as the current literature suggests that women who stopped

smoking during the first trimester had similar risks of HDP as those who were non-smokers. [70,71] To adjust for the quality of prenatal care, we used the Revised Graduated Prenatal Care Index (R-GINDEX). The R-GINDEX compares the reported number of prenatal care visits to how many visits a pregnant person is *expected* to have, based on ACOG guidelines, adjusting for the gestational age at the first prenatal appointment and length of the pregnancy. This metric is classified into six categories – inadequate, intermediate, adequate, intensive, no care, and missing. [72] Inadequate care was defined as receiving less than 50% of expected visits, intermediate between 50-79% visits, adequate 80%-109%, and intensive (previously called adequate plus) was greater than 110%. [73]

Each record was geocoded using the ESRI address coder (ESRI, Redlands, CA). Records with coordinates corresponding to the "rooftop" or a statement were classified as "precisely geocoded." Coordinates that corresponded to the midpoint of a street, ZIP code, or city were considered imprecise. [45] For further details on the geocoding methods and precision of this dataset, see Chapter 2. Using standard geocoding convention, we considered counties with more than 85% of addresses geocoded as high precision areas; otherwise, they were considered low precision areas. [45] To create the county-level dataset, we used the individual records to characterize the yearly county-level prevalence of mothers of advanced maternal age (% $\geq$ 34 years old), race (% Black), ethnicity (% Hispanic), educational attainment (% completed college), marital status (% married), pre-existing diabetes (%), maternal obesity, maternal smoking throughout pregnancy (%), and stillbirths (%). Prevalence estimates were then classified into quartiles using PROC RANK. Rurality status, geocoding precision, and Appalachian designation was determined at the county level. [66] [74]

3.3.3 Statistical Analysis

Summary Statistics

We summarized all covariates of interest as counts and percentages. To calculate the weighted average eHDP cases within each covariate level, we used the LS MEANS option within the PROC GENMOD with a negative binomial distribution and a log link.

Bivariate and multivariate models

To screen for multicollinearity, we calculated Spearman's rank pairwise correlation. No two variables had a rho greater than 0.6.

For both the bivariate and multivariate models, we fit a fixed-effects negative binomial regression model for longitudinal data with an autoregressive (AR) correlation structure, offset with the natural log of the number of births in each county and year using PROC GENMOD. Time was treated as a categorical variable in all models. Fixed effects allowed us to adjust for the similarity within one county over time. The negative binomial model was selected because the mean and variance structure assumption was violated for the Poisson model. The AR correlation structure was chosen because it allows for a stronger correlation between temporally closer times, and the strength of association is assumed to reduce as distance among time points increases.

For the bivariate model, initially, we used the negative binomial described above. We assessed each covariate interacted with time (categorical) to explore eHDP prevalence in relation to the changes in each covariate over time. However, none of the interactions were statistically significant; therefore, we reported the non-interacted results for brevity.

For the final model, variables identified in the literature as important covariates were included in the base model |maternal age $\geq$ 34 years, (%), race (Black %), maternal obesity (%), and smoking throughout pregnancy (%)|. All other covariates were removed with backward elimination. Variables that were statistically significant or changed the estimates of statistically significant covariates by more than 15% were retained in the model. The final model included maternal age $\geq$ 34 years, (%), race (Black %), educational attainment (% completed college), marriage (%), maternal obesity (%), smoking throughout pregnancy (%), and the Appalachian region.

All analyses were conducted using used SAS v 9.4 (SAS Corp., Cary, NC)

Mapping and temporal trend assessment

To explore geographic patterns of eHDP and detect and evaluate the statistical significance of any identified clusters, we performed unadjusted retrospective space-time cluster analyses using the SaTScan (v 9.5) software.

SaTScan™ is a trademark of Martin Kulldorff. The SaTScan™ software was developed under the joint auspices of (i) Martin Kulldorff, (ii) the National Cancer Institute, and (iii) Farzad Mostashari of the New York City Department of Health and Mental Hygiene.

Briefly, this method delineates several overlapping cylinders of varied sizes and widths over the study area to identify possible clusters of cases in space and time. [75] For this study, each cylinder was centered on a point in a regular 5-mile grid and could encompass various surrounding counties. Generally, each cylinder's radius corresponds to geographic distance, and the height corresponds to time. In our study, we focused only on high-prevalence clusters of at least two cases. Maximum spatial cluster size was initially set to 30% of the study area population, as this would capture large cities, such as Louisville. However, no clusters were identified in urban areas, and the identified clusters were too large to be useful (e.g., 42 counties). Therefore, the maximum size of the spatial clusters was gradually reduced by five percent until the number of counties identified was narrow enough in scope to identify potential areas for intervention. The final spatial cluster size was 10% of the covariate adjusted population at risk. We also assessed purely spatial clusters or those elevated for the entire study period. Under the null hypothesis, we assumed that cases were Poisson distributed and risk was constant over space and time. The alternative hypothesis was that the risk would be higher inside the cluster than outside the cluster.

We created choropleth maps using QGIS (Madeira v 3.4) to display identified clusters and visualize the average prevalence of eHDP, marriage, maternal obesity, maternal smoking throughout pregnancy for each county over the duration of the study. We used the Jenks method to determine categories for choropleth maps. [76]

We used a general linear estimation (GLM) model with a Poisson distribution and a log link to obtain yearly estimates of eHDP and the average annual percent change (AAPC). Significant covariates (maternal obesity, smoking throughout pregnancy, marriage, and eHDP) were assessed for significant inflection points using Joinpoint software. [77]

3.4 Results

In this retrospective ecological study, we observed 1,936 cases of eHDP among 212,544 births (9.1 cases per 1,000 births) in Kentucky from 2008 through 2017. Table 3.1 displays the marginal means of eHDP for each subgroup. Counties with the highest prevalence of Black mothers had the lowest average number of eHDP cases (8.1, 95% CI: 7.4, 9.0), as did the areas with the lowest obesity prevalence (7.9, 95% CI: 7.2, 8.5). The Appalachian region had one of the highest marginal means with 11.1 eHDP cases on average per 1,000 births (95% CI: 10.2, 12.0).

Table 3.1 Number of counties in each demographic group, and average eHDP prevalence by demographic group

	N (%)	County level average eHDP (95% CI)
Mothers ≥34 years old (%)		
<2.5	299 (24.9)	10.75 (9.45, 12.50)
2.5 to 4.2	300 (25.0)	10.20 (9.20, 11.50)
4.2 to 6.1	302 (25.2)	9.95 (9.00, 11.00)
>6.1	299 (24.9)	8.35 (7.70, 9.00)
Race (Black %)		
0	436 (36.3)	10.45 (9.30, 11.50)
0.1-1.7	193 (16.1)	10.05 (8.95, 11.50)
1.7 to 3.4	188 (15.7)	9.80 (8.70, 11.00)
3.4 to 6.2	189 (15.8)	10.15 (9.05, 11.50)
>6.2	194 (16.2)	8.05 (7.35, 9.00)
Ethnicity (% Hispanic)		
0	386 (32.2)	11.45 (10.15, 13.00)
0.1-1.7	205 (17.1)	9.70 (8.65, 11.00)
1.7 to 2.8	198 (16.5)	10.05 (8.95, 11.50)
2.8 to 4.5	209 (17.4)	9.40 (8.40, 10.50)
>4.5	202 (16.8)	8.10 (7.40, 9.00)
Educational attainment (% Completed college)		
<18.0	299 (24.9)	10.00 (8.75, 11.50)
18.0 to 23.4	301 (25.1)	10.55 (9.45, 12.00)
23.4 to 30.2	301 (25.1)	9.55 (8.60, 10.50)
>30.2	299 (24.9)	8.80 (8.15, 9.50)
Marriage (%)		
<43.6	300 (25.0)	8.95 (7.75, 10.50)
43.6 to 48.6	300 (25.0)	9.70 (8.75, 11.00)

48.6 to 53.8	298 (24.8)	9.30 (8.45, 10.00)
>53.8	302 (25.2)	9.80 (8.95, 10.50)
Pre-existing diabetes (%)		
0	551 (45.9)	9.40 (8.45, 10.50)
0.1-0.8	159 (13.3)	8.40 (7.65, 9.00)
0.8 to 1.2	163 (13.6)	9.60 (8.60, 10.50)
1.2 to 1.9	162 (13.5)	10.75 (9.40, 12.50)
>1.9	165 (13.8)	11.80 (10.05, 14.00)
Maternal obesity (%)		
<22.6	298 (24.8)	7.90 (7.20, 8.50)
22.6 to 26.8	301 (25.1)	9.85 (8.95, 11.00)
26.8 to 31.6	301 (25.1)	10.50 (9.50, 11.50)
>31.6	300 (25.0)	10.85 (9.65, 12.00)
Maternal Smoking throughout pregnancy (%)		
<13.2	301 (25.1)	8.50 (7.85, 9.00)
13.2 to 17.8	298 (24.8)	10.10 (9.15, 11.00)
17.8 to 23.0	301 (25.1)	10.40 (9.30, 11.50)
>23.0	300 (25.0)	10.15 (8.95, 11.50)
Prenatal care (% Inadequate*)		
<5.7	297 (24.8)	10.30 (9.20, 11.50)
5.7 to 8.6	304 (25.3)	9.10 (8.35, 10.00)
8.6 to 12.4	300 (25.0)	9.30 (8.40, 10.00)
12.4 to 35.7	299 (24.9)	9.80 (8.70, 11.00)
Stillbirths (%)		
0	647 (53.9)	9.50 (8.70, 10.50)
0.1-0.7	181 (15.1)	8.90 (8.20, 9.50)
0.7 to 1.5	191 (15.9)	9.85 (8.80, 11.00)
>1.5	181 (15.1)	11.35 (9.65, 13.50)
Rurality		
Metro/Non-Metro	850 (70.8)	10.60 (9.90, 11.50)
Rural	350 (29.2)	8.35 (7.80, 9.00)
Geocoding precision		
<85% geocoded precisely	468 (39.0)	10.35 (9.40, 11.50)
>85% geocoded precisely	732 (61.0)	9.15 (8.60, 10.00)
Appalachian Region		
Non-Appalachian	660 (55.0)	8.70 (8.15, 9.50)
Appalachian	540 (45.0)	11.10 (10.20, 12.00)

Year		
2008	120 (10.0)	8.10 (6.85, 9.50)
2009	120 (10.0)	8.45 (7.20, 10.00)
2010	120 (10.0)	9.10 (7.75, 10.50)
2011	120 (10.0)	8.50 (7.20, 10.00)
2012	120 (10.0)	9.50 (8.10, 11.00)
2013	120 (10.0)	10.95 (9.45, 12.50)
2014	120 (10.0)	9.60 (8.20, 11.50)
2015	120 (10.0)	10.35 (8.85, 12.00)
2016	120 (10.0)	10.10 (8.60, 12.00)
2017	120 (10.0)	10.45 (8.90, 12.50)

*Inadequate care: Receiving less than 50% of expected prenatal visits, Maternal obesity: BMI >30 kg/m^2

Table 3.2 displays the unadjusted and adjusted models. In the unadjusted model, we observed that counties with the lowest proportion of mothers $\geq$34 years old had a 26% higher prevalence of eHDP (RR=1.26, 95% CI: 1.08, 1.52) than counties with the highest proportion of mothers $\geq$ 34 years old. Also of note in the unadjusted model was the 33% reduction in eHDP prevalence in counties with the highest percentage of Black mothers, compared to counties with no Black mothers.

In the final model, adjusted for mothers $\geq$ 34 years old (%), race (Black %), marriage (%), maternal obesity (%), maternal smoking throughout pregnancy (%), the Appalachian region, and year, we observed that low proportions of mothers $\geq$ 34 years old, and higher proportions of maternal obesity (%) and marriage (%) were associated with an increased prevalence of eHDP. In the unadjusted model, race showed a statistically significant decrease in eHDP prevalence; however, in the adjusted model, this relationship shifted to a non-significant *increase* in prevalence compared to counties with no Black mothers (RR=1.04, 95% CI: 0.86, 1.26). The proportion of married mothers, insignificant in the unadjusted model, increased the prevalence of eHDP by 38% in the adjusted model (95% CI: 1.16, 1.64). The relative risk of eHDP in the Appalachian region remained relatively unchanged following covariate adjustment (RR:1.22, 95% CI:1.11, 1.44, aRR:1.18, 95% CI:1.02, 1.42).

Table 3.2 Unadjusted and adjusted prevalence of eHDP by demographic category

	PR (95% CI)	p-value	aPR (95% CI)	p-value
Mother ≥34 years old (%)				
<2.5	1.20 (1.08, 1.52)	<0.01	1.26 (1.04, 1.50)	0.01
2.5 to 4.2	1.21 (1.00, 1.43)	0.02	1.16 (1.00, 1.38)	0.05
4.2 to 6.1	1.18 (1.02, 1.37)	0.02	1.14 (1.00, 1.30)	0.06
>6.1	Reference		Reference	
Race (Black %)				
>6.2	0.77 (0.63, 0.93)	<0.01	1.04 (0.86, 1.26)	0.62
3.4 to 6.2	0.96 (0.80, 1.15)	0.72	1.18 (1.00, 1.38)	0.05
1.7 to 3.4	0.94 (0.77, 1.14)	0.55	1.10 (0.92, 1.30)	0.30
0.1-1.7	0.96 (0.81, 1.13)	0.68	1.02 (0.86, 1.18)	0.88
0	Reference		Reference	
Educational attainment (% completed college)				
>30.2	0.88 (0.73, 1.05)	0.18	1.02 (0.82, 1.28)	0.81
23.4 to 30.2	0.95 (0.79, 1.13)	0.57	0.92 (0.76, 1.10)	0.38
18.0 to 23.4	1.05 (0.91, 1.21)	0.47	1.00 (0.86, 1.18)	0.95
<18.0	Reference		Reference	
Marriage (%)				
>53.8	1.10 (0.90, 1.30)	0.23	1.38 (1.16, 1.64)	<0.01
48.6 to 53.8	1.04 (0.88, 1.24)	0.58	1.18 (1.02, 1.40)	0.03
43.6 to 48.6	1.09 (0.90, 1.30)	0.35	1.16 (0.98, 1.36)	0.08
<43.6	Reference		Reference	
Maternal obesity (%)				
>31.6	1.37 (1.13, 1.65)	<0.01	1.20 (1.00, 1.44)	0.05
26.8 to 31.6	1.32 (1.12, 1.50)	<0.01	1.14 (0.98, 1.36)	0.11
22.6 to 26.8	1.24 (1.07, 1.43)	<0.01	1.14 (1.00, 1.32)	0.07
<22.6	Reference		Reference	
Maternal smoking throughout pregnancy(%)				
>23.1	1.19 (1.00, 1.41)	0.04	1.16 (0.94, 1.42)	0.15
17.9 to 23.1	1.22 (1.04, 1.42)	<0.01	1.20 (1.02, 1.42)	0.03
13.2 to 17.9	1.18 (0.99, 1.42)	0.05	1.16 (1.02, 1.32)	0.02
<13.2	Reference		Reference	
Appalachian Region				
Appalachian	1.20 (1.11, 1.44)	<0.01	1.18 (1.02, 1.36)	0.02
Non-Appalachian	Reference		Reference	

Year				
2017	1.29 (1.06, 1.57)	0.01	1.34 (1.08, 1.66)	0.01
2016	1.24 (1.04, 1.48)	0.01	1.36 (1.14, 1.64)	<0.01
2015	1.28 (1.06, 1.54)	0.01	1.34 (1.12, 1.60)	<0.01
2014	1.10 (0.97, 1.45)	0.09	1.28 (1.04, 1.58)	0.02
2013	1.35 (1.11, 1.66)	0.01	1.46 (1.22, 1.74)	<0.01
2012	1.17 (0.93, 1.46)	0.16	1.22 (1.00, 1.50)	0.04
2011	1.05 (0.84, 1.31)	0.64	1.10 (0.90, 1.34)	0.39
2010	1.12 (0.93, 1.37)	0.22	1.16 (0.98, 1.38)	0.08
2009	1.04 (0.85, 1.27)	0.66	1.08 (0.90, 1.32)	0.39
2008	Reference		Reference	

Maternal obesity: BMI >30 kg/m^2

Three clusters were identified in the spatial analyses: two in Eastern Kentucky (part of the Appalachian region) and one in Western Kentucky (Figure 3.1). The largest cluster, located in Eastern Kentucky, was comprised of 14 counties (Table 3.1). Three counties in the largest cluster (Breathitt, Harlan, and Letcher) had more than 15 eHDP cases per 1,000 births throughout the study period (2008-2017). The smallest cluster of eHDP in Appalachia, comprised of two counties, was limited to only one year, 2012. The third cluster, elevated from 2013-2017, approached significance (p=0.1) and comprised four counties in Western Kentucky.

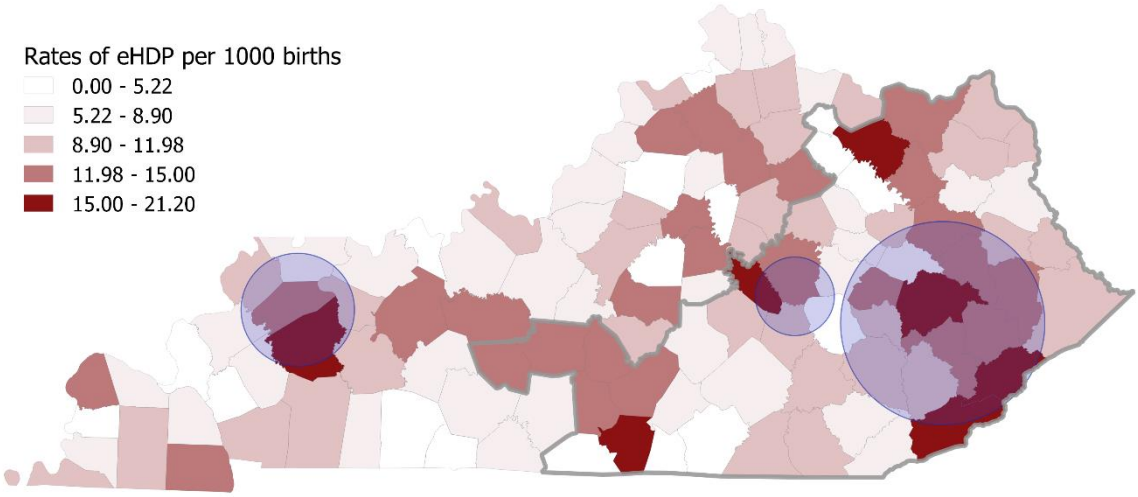


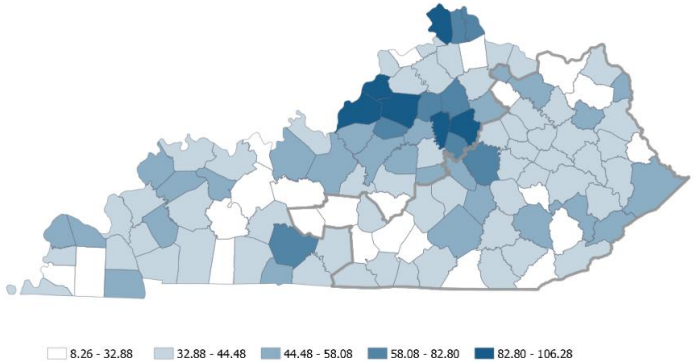
Figure 3.1 Choropleth map of the county-level prevalence of eHDP in Kentucky per 1000 births and high-rate clusters of eHDP prevalence, 2008-2017

Table 3.3 Details of identified clusters of elevated eHDP prevalence, 2008-2017

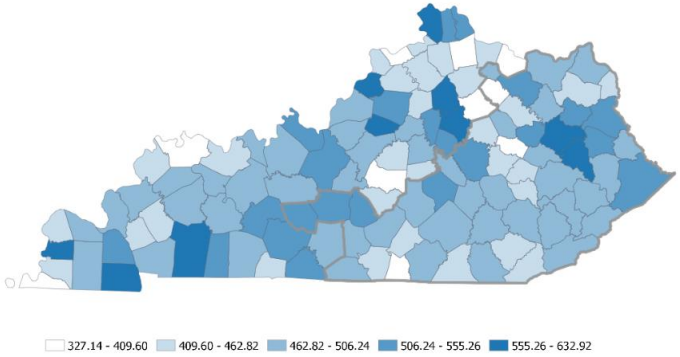
Cluster description Counties	**Years**	**In Cluster** **Case | Pop**	**Out Cluster** **Case | Pop**	**RR**	**P-value**
1. Large cluster – Eastern KY: Breathitt, Perry, Knott, Magoffin, Leslie, Owsley, Wolfe, Lee, Clay, Floyd, Morgan, Letcher, and Harlan	2008-2017	148| 1083	1784 | 18339	1.54	0.03
2. Small cluster – Central KY: Madison and Garrard	2012	18 | 466	176 | 20830	3.81	0.05
3. Small cluster in Western KY: Webster, Hopkins, McLean, Union	2013-2017	38| 388	965 | 21500	2.28	0.10

N: Number; RR: Relative Risk, Pop: Population

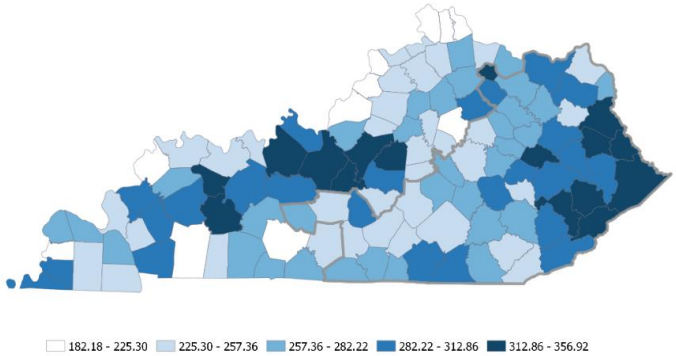
A. Births to women ≥34 years old (per 1000 births)



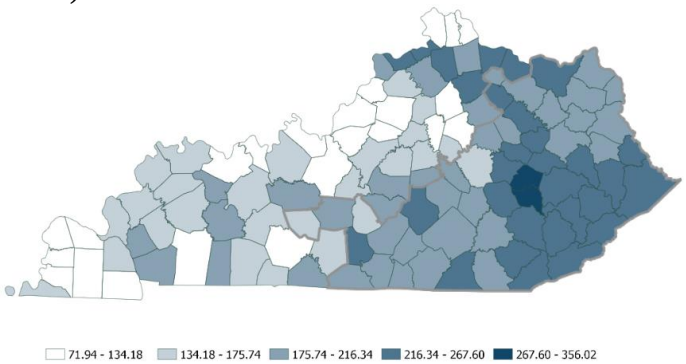
B. Married mothers (per 1000 births)



c. Maternal Obesity (per 1000 births)



d. Maternal smoking throughout pregnancy (per 1000 births)



Maternal obesity: BMI >30 kg/m^2

Figure 3.2 Choropleth maps of county-level proportions of marriage, maternal obesity, and maternal smoking per 1000 births in Kentucky, 2008-2017

To assess spatial trends in the covariates of interest, we created choropleth maps displaying the prevalence of maternal obesity, marriage, maternal age ≥34, and maternal smoking per 1000 births (Figure 3.2). It is important to note that the maps summarize *the entire study period* and are presented for illustrative purposes. We statistically assessed the residuals for each county and year with Moran's I with GeoDa (v 1.18, December 2020). There was no indication of patterns of poor model fit. [78]

There was a non-statistically significant increase in the prevalence of eHDP (AAPC:2.8, 95% CI: -4.3, 10.5) over the study period (Table 3.4). Both prevalence of maternal obesity (AAPC:2.3, 95% CI: 0.94, 3.7) and births to mothers older than 34 years (AAPC:3.4, 95% CI: 0.66, 6.3) increased over the study period. Maternal smoking decreased by almost 6% (AAPC:-5.8%, 95% CI: -7.5, -4.1). Upon visual inspection, there appeared to be a shift in the prevalence of obesity and smoking in 2012; however, further investigation using Joinpoint to assess inflection points yielded non-significant results.

Table 3.4 Prevalence of eHDP, marriage, maternal obesity, and maternal smoking per 1000 births in Kentucky, 2008-2017

Year	eHDP	Married	Maternal obesity	Smoking during pregnancy	Maternal age ≥34 years
2008	7.68	504.04	232.00	182.90	56.72
2009	8.22	501.46	232.98	164.66	58.70
2010	8.68	503.62	230.24	152.18	54.92
2011	8.12	499.34	240.48	149.52	58.52
2012	9.06	504.98	243.56	153.38	61.22
2013	10.76	504.70	243.36	141.40	61.46
2014	9.50	513.54	255.46	127.18	64.46
2015	9.84	513.50	264.78	118.46	70.52
2016	9.76	513.86	269.34	110.90	71.86
2017	9.88	521.16	284.52	99.56	74.94
AAPC (95% CI)	2.84 (-4.26, 10.46)	0.40 (-0.56, 1.36)	2.32 (0.94, 3.74)‡	-5.80 (-7.52, -4.06)‡	3.44 (0.66, 6.28) †

eHDP: Early-onset HDP (HDP onset <34 weeks), Maternal obesity: BMI >30 kg/m^2

AAPC: Annual Average percent change CI: Confidence Interval, † p-value <0.05, ‡ p-value <0.01

3.5 Discussion

This retrospective ecological study sought to characterize the prevalence of eHDP, identify significant county-level covariates associated with increased eHDP prevalence, and identify potential geospatial patterns of eHDP in Kentucky. This study demonstrated a low county-level prevalence of primiparous mothers ≥ 34 years old, and high county-level prevalences of maternal obesity and smoking throughout pregnancy were associated with increased prevalence of eHDP. Additionally, we detected two statistically significant clusters of eHDP in the Appalachian region of the state and one cluster approaching statistical significance in Western Kentucky.

The prevalence of eHDP in Kentucky from 2008-2017 was approximately 9.2 cases per 1,000 births. Although there are no nationwide estimates of eHDP prevalence, a study using birth records from Washington reported 3.8 eHDP per 1,000 births – slightly less than half of Kentucky rates. [57] These findings may reflect the general elevation of obesity and pre-existing diabetes in Kentucky relative to Washington; however, further study is needed, as other factors, such as environmental regulations and predominant industry, vary between these states. [79,80]

We found that county-level prevalence of maternal age ≥ 34, maternal obesity, maternal smoking throughout pregnancy, and marriage was significantly associated with eHDP; however, contrary to existing literature, increased prevalence of marriage and maternal smoking was associated with increased rates of eHDP. [81] This finding may be due to an ecological fallacy – an inappropriate attribution of individual risk based on ecological (in this case county-level) data. Or, our findings could be an indication that the risks of eHDP may differ from late-onset HDP or other subsets of HDP. [59]

The detection of two statistically significant clusters of eHDP within the Appalachian region, an area that has historically had elevated rates of obesity and smoking, is also noteworthy. This study found that within the Appalachian region, eighteen counties had a maternal obesity prevalence of 30% or greater, twelve counties reported 25% or more mothers engaged in smoking throughout their pregnancy, and six of the seven counties with an eHDP prevalence greater than 15 cases per 1000 births

were located in the Appalachian region. This, too, maybe the result of ecological fallacy or reflect the coincidence of obesity and high-risk health behaviors. Given the cross-sectional nature of this study, we can not assess the causal associations between maternal obesity, smoking, and eHDP; however, further assessment of these relationships is needed, as previous studies of individual risks have found smoking to be protective.

This study found that, over the study period, there was a 3% increase in eHDP prevalence. Wallis and colleagues reported that PE, a subset of HDP, increased by 29.4%, and GH by 30.6% over 17 years (1987-2004). The rapid increase in HDP experienced in the Wallis study may reflect substantial changes within the population over the 17 years or could be due to the inclusion of all HDP cases, unlike in this study that focused only on early-onset disease. However, it is important to note that there were two changes in the definition that case definitions of gestational hypertension and pre-eclampsia changed over the study period, which may have changed diagnoses and impacted rate calculations. [8]

The proportion of first-time mothers $\geq$34 years old increased over the study period, particularly in urban areas such as Louisville, Northern Kentucky, and Fayette County. Madison County, just south of Fayette, also had a high prevalence of primiparous mothers $\geq$34 years old – an anomaly in the Appalachia region. However, this may be due to the presence of two post-secondary institutions within this relatively small county. Existing research has suggested that obtaining a college education or higher is associated with delayed fertility; however, further research is needed. [82] In this study, we found that the largest high-prevalence cluster of eHDP largely comprised of counties with the lowest proportion of mothers $\geq$34 years of age. In other words, within Kentucky, women <34 could be driving the increase in eHDP. Although this finding may be due to the limitations inherent within ecological studies, notably ecological fallacy, and poor reproductive health literacy among young adults, especially in high-poverty regions, may also contribute to these findings.

Unanticipated findings of this study were the 11-year trend of increasing eHDP prevalence in areas with a high proportion of married women. Marriage is often identified as protective against HDP, as long-term sexual partners and non-barrier methods of

contraception are associated with reduced risk of HDP. [83,84] The increased prevalence of eHDP found in communities with a high prevalence of maternal smoking was also surprising, as smoking throughout pregnancy has previously been associated with a decreased risk. This finding may be due to our exclusive focus on women who gave birth early; however, findings of this type require further study at the individual level.

We also identified that maternal smoking prevalence has decreased by almost 6% and obesity increased by over 2% over the study period. Further work evaluating the larger impact of these demographic changes and what may be a contributing factor to each covariate's positive and negative growth is needed.

Limitations and Strengths

Given some of our unexpected findings, unmeasured confounding from factors such as maternal stress during pregnancy, environmental or occupational exposures, and other environmental characteristics may be influencing our estimates. [15,85-89] However, as an ecological study, there is the risk of ecological fallacy, or the trends we see at the county level may not be representative of or exist in the same manner at the individual level.

This study used only birth record data for maternal demographic information, pregnancy characteristics, and birth outcomes, which has been shown to have underreporting biases with other pregnancy conditions. [90] The collection of birth records have also been found to vary among Kentucky hospitals. [91] We believe any biases introduced due to clerical error are non-differential as there is no indication that misclassification occurred based on any covariates of concern. [91,92] Additional limitations in this study include the case definition. We were also unable to determine the gestational age of disease onset from the birth record. To ensure we had identified cases of early-onset HDP, we used births before 35 weeks gestational age as the cut point. However, women could have developed eHDP but given birth at or after 35 weeks.

Further, the certificates of live and stillbirth do not distinguish between GH, PE, and HELLP syndrome and are mutually exclusive with a diagnosis of chronic hypertension as pre-existing hypertension is a noted risk factor for HDP. [65] However, the diagnostic criteria for PE and GH were modified in 2013, as these are not individually

distinguishable on the birth record; we do not expect it to have overly impacted rates. However, changes in procedure can lead to the over or underreporting of cases. [5,8] There may be interviewer and recall biases due to questions that might be perceived as stigmatizing – such as self-reported pre-pregnancy weight, marital status, number of cigarettes smoked throughout the pregnancy.

Geocoding maternal addresses provided on the birth certificates is another source of potential bias in this study, as the maternal county of residence was determined from the mother's residential address. Previous research has shown that Kentucky birth records in rural areas geocode less precisely than their urban counterparts (Chapter 2). However, given the overall precision of the records and the spatial unit of analysis being at the county level, we believe that the impact of geocoding imprecision on cluster identification to be limited. However, changes in residences, which are not reported on the birth certificate, may have led to non-differential misclassification of women who changed residence during pregnancy. [93]

There are some notable strengths of this study. The certificates of live and stillbirth are established administrative data collection forms that undergo routine quality control. Furthermore, the form's questions remained relatively unchanged throughout the study period. The form captures important demographic information (maternal age, race, ethnicity, and education) and pregnancy characteristics (gestation, prenatal care, pre-pregnancy weight, height, and gestational age at birth) to better adjust for confounders. [94-96] We had sufficient study power to detect a statistically significant spatio-temporal cluster of eHDP, an infrequent pregnancy complication.

Although research on individuals have suggested that smoking is protective, this study shows that on a county level, smoking may be indicative of other health conditions unaccounted for in the birth record which could have increased prevalence of eHDP. [97] Furthermore, this study confirms that the prevalence of eHDP and known risk factors, notably obesity, are increasing in Kentucky, and targeted health intervention is needed to improve maternal and fetal health. Lastly, this study has shown that there are clear spatial patterns in the incidence of eHDP incidence. Further research exploring explanations for high-rate clusters, such as air and water quality, occupational exposures, and the impact

of community amenities that encourage physical activity (green spaces and access to healthy foods), will further help identify potential interventions that could alleviate the disease burden.

CHAPTER 4. A CROSS-SECTIONAL EXAMINATION OF THE ASSOCIATION OF ENVIRONMENTAL TOXICITY WITH EARLY-ONSET HYPERTENSIVE DISORDERS OF PREGNANCY IN KENTUCKY 2008-2017

4.1 Abstract

This exploratory study merged geocoded residential information from birth records (2008-2017) with aerosol toxicity concentration estimates from the Risk Screening Environmental Indicators (RSEI) model (2007-2017) to assess geospatial patterns in industrial emissions of toxic metals (arsenic, cadmium chromium, lead, and mercury) and to evaluate the prevalence of early-onset hypertensive disorders of pregnancy (eHDP) in association with exposures to this set of metals. Four distinct classes of toxic metal exposure were identified using a latent class analysis. Women classified as having a high probability of exposure to both lead (Pb) and chromium (Cr) had a statistically significantly greater prevalence of eHDP (aPR=1.22, 95% CI: 1.04, 1.44), relative to those with low or no exposure. Our findings contribute to the emerging literature on the association of metal exposures with pregnancy outcomes.

4.2 Introduction

Exposure to environmental toxicants has been shown to increase the risk of respiratory and cardiovascular disease, breast cancers, and poor pregnancy outcomes such as hypertensive disorders of pregnancy (HDP). [98-100] Hypertensive disorders of pregnancy, generally considered progressive, impacts 8-10% of pregnancies in the US each year, and are among the leading causes of morbidity and mortality in mothers and infants. Subsets of HDP include gestational hypertension (GH), pre-eclampsia (PE), and eclampsia. [1,2] Short-term adverse events of HDP can include preterm birth, stroke, and renal failure. Even after the resolution of the pregnancy, women who experience more severe HDP, such as PE, are at an increased risk of hypertension, stroke, metabolic disease, and HDP in subsequent pregnancies. [3,4] Pharmacologic interventions and treatment are extremely limited. Thus far, only low-dose aspirin has been shown to be minimally effective at reducing the risk of HDP if taken beginning at 12 weeks gestation; use after symptom onset has not been shown to be effective. [101] Risk factors for HDP include primiparity, maternal age, obesity, race, and use of infertility treatment. [6]

Preliminary research suggests the presence of heavy metals in the maternal system may also be associated with an increased risk of HDP; however, some findings have been inconclusive, notably for As and Hg. [14-21] These contradictory findings may indicate no relationship between HDPs and specific metal exposure; however, given substantial variation in the frequency and type of the biological samples used among studies, further study is needed.

Also of concern with many HDP studies is their ability to generalize to US maternal populations. Although the underlying biological mechanisms resulting in disease would be consistent regardless of geography, maternal diet, environmental regulations, and health infrastructure could vary enough to impact both volume/types of exposures and the likelihood of disease. [102]

Inconsistencies in sampling methodologies have also plagued studies assessing environmental concentrations of metal exposure in air or water, leading to challenges in synthesizing results. [34] One potential option to address this limitation is to use the Environmental Protection Agency's Risk Screening Environmental Indicators model. The Risk Screening Environmental Indicators (RSEI) approach uses emissions data reported to the Toxic Release Inventory (TRI) program to characterize yearly ambient aerosol concentrations of individual chemicals of concern across the U.S., adjusted for physicochemical properties and site characteristics (such as stack height, when available). [35] These data are attractive, particularly in pilot studies, as they are easily accessible, include over 700 chemicals tracked by the EPA, are weighed for their overall toxic effects on human health and can be linked to administrative boundaries, such as ZIP codes, census tracts, or micro-block groups. [35]

To assess the utility of RSEI data, we selected birth records from Kentucky, a state with a high prevalence of eHDP risk factors (obesity, pre-existing diabetes) [103] and a high prevalence of smoking, which has generally been found to be protective. [104] The state also hosts large industrial complexes in both urban and rural areas, making it well suited to explore the relationship between environmental metal exposures and eHDP. [105] In this study, we had four aims: 1) examine the distribution of emissions of chemicals of concern across the state 2) identify patterns of exposure to industrial metal emissions and

describe the sociodemographic characteristics of mothers in these areas, 3) evaluate the impact of environmental exposures to industrial metal emissions, adjusting for sociodemographic factors on risk eHDP and 4) identify areas in the state that have a high prevalence of individuals with eHDP. We hypothesize that women living in areas with an overlap in elevated exposures to As, Cd, and Pb during pregnancy would have a higher probability of eHDP than those living in areas with singular exposures. We also expected that women who lived in areas with elevated Se or Zn concentration would be less likely to have an eHDP diagnosis, as Se and Zn have been shown to be protective against the effects of cadmium toxicity. [106]

4.3 Methods

The Kentucky Cabinet for Health and Family Services IRB and the University of Kentucky IRB reviewed and approved this cross-sectional study. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were used as a template for reporting this study.

4.3.1 Study Population

The Kentucky Department of Vital Statistics provided 557,751 individual records for all live (n=553,476) and stillbirths (n=3,268) to self-identified Kentucky residents from January 1, 2008, through December 31, 2017. All covariates assessed for model inclusion were present in both forms. [64] Records were excluded if the mother had chronic hypertension, as it is mutually exclusive with HDP on birth records (n=10,752), non-primiparous (n=327,459), a multifetal pregnancy (n=5,206), to a mother younger than 11 or older than 50 years old (n=215), delivered before 20 weeks gestation or after 45 weeks (n=565), the sex of the child was not known (n=20) the record did not geocode (n=3) or geocoded outside of the state (n=473), leaving 212,051 for analysis (1247 stillbirth and 210,804 live births) for analysis. [65] Early-onset hypertension (eHDP) was defined as giving birth before 34 weeks and being positive for gestational hypertension, which includes diagnoses of gestational hypertension, pre-eclampsia, or HELLP (Hemolysis, ELevated Liver enzymes, Low Platelet count), on the birth record. [65] Individual-level data were assessed for patterns in missingness using PROC MI in SAS.

4.3.2 Outcome and Exposure Ascertainment

Rural-Urban Continuum Codes (RUCC), a metric of population density, were obtained from the United States Department of Agriculture. The Area Deprivation Index (ADI) was used as a metric for the economic conditions of a given census tract. [107]

The ADI dataset incorporates American Community Survey (ACS) data on income, housing, educational, and employment data into principal component analysis to derive a score that is then standardized across the state. Higher scores correspond to higher levels of economic distress. [48] For this study, the 2019 Kentucky-specific census micro-block group ADI data were linked to the birth records by the geocoded census micro-block group. The ADI scores were dichotomized into upper quintiles (ADI of 9-10) and lower (ADI 1-8) of neighborhood poverty. [107] Cartographic boundary files for county and census micro-block groups were obtained from the United States Census Bureau. [67] Appalachian status, defined by the Appalachian Regional Commission (ARC), was based on the geocoded maternal county of residence. [68]

To assess environmental exposure, disaggregated census micro-block group (CMBG) air emissions data were requested and obtained from the Environmental Protection Agency's (EPA) Risk Screening Environmental Indicators (RSEI) program for the years 2007 through 2017. [108] Briefly, RSEI uses the yearly chemical-specific stack and fugitive aerosolized emissions reported to the Toxic Release Inventory (TRI) by federal and mandated facilities. Facilities are required to report if: 1) they have ten or more full-time employees, 2) the industry is in a required sector (such as mining), or is a federal facility and 3) manufactures, processes, or uses TRI-listed chemicals and the production, use, or transfer of a chemical amount exceeds the threshold set for a chemical in a given category. [109]

The PROC RANK procedure in SAS was used to identify and classify the highest quintiles for each exposure. As few CMBG had an estimated toxicity concentration of cadmium, records that had *any* cadmium exposure in their respective CMBG were characterized as such in a binary fashion. The yearly RSEI data were linked to birth records by CMBG using the residential address and year the majority (>50%) of the pregnancy period took place, including a 12-week preconception period to account for

pre-pregnancy exposures. [110,111] Therefore, although birth records were from 2008-2017, we used exposure data from 2007-2017.

4.3.3 Data cleaning and dataset preparation

Individual-level covariates were obtained from the birth certificate records. Maternal age was calculated by subtracting the mother and infant date of birth and rounding down to the nearest year, then categorized into five groups (>20, 21-24, 25-28, 29-34, >35). Maternal race was collapsed into three categories (Black, Other, White). Maternal Body Mass Index (BMI) was calculated using self-reported height and pre-pregnancy weight and categorized into: underweight/normal (<25 kg/m^2), overweight (25-30 kg/m^2), and obese (>30 kg/m^2). [69] Current literature suggests that women who quit smoking during their first trimester have an equivalent risk of HDP as women who do not smoke. Therefore, those who reported no smoking throughout their pregnancy or reported no cigarette use after the second trimester were considered non-smokers. Otherwise, women were classified as smokers. [70] Other covariates captured on the birth record included maternal ethnicity (Hispanic/non-Hispanic), education (less than high school, high school, some college, and college degree), marital status (yes/no, or not stated), and pre-existing diabetes (yes/no, or not stated).

Each record was geocoded using an ESRI address coder (ESRI, Redlands, CA). This process provides geographic coordinates, the precision of coordinates, and local administrative boundaries (county and census tract information) for each address. Precisely geocoded addresses were those that were identifiable at the rooftop or street segment. Addresses that geocoded to the centroid of a street, city, or ZIP code were considered imprecise. Further details are geocoding methodology are provided in Chapter 2. A summary of the geocoding precision of the records is reported in Appendix B.

4.3.4 Statistical analysis

Spatial Analysis

To assess geospatial patterns of disease, SaTScan software was used to conduct a retrospective spatiotemporal scan statistic (Bernoulli model) with a circular scan window to detect clusters of high eHDP rates in Kentucky over the time period. [75] This method identifies candidate clusters using overlapping cylinders of increasing heights and

diameters representing time and spatial dimensions, respectively until a user-defined maximum population (10%) and temporal inclusion (5 years) restriction is reached. The maximum population is defined by the number of cases and non-cases in the input file. Using the likelihood ratio test, SaTScan compares the number of observed and expected cases within a candidate cluster to the area outside the cluster, adjusting for the underlying population density. The Monte Carlo method was used to estimate the p-value.

Both purely spatial candidate clusters, or those that spanned the entire study period, and clusters that encompassed 50% or less of the study period were assessed. [75] To identify high-rate clusters of eHDP, the maximum cluster size was restricted to 10% of the population after confirming that larger population centers such as Lexington and Louisville had no clusters, as population size restrictions would effectively exclude them from reported results.

SaTScan™ is a trademark of Martin Kulldorff. The SaTScanTM software was developed under the joint auspices of (i) Martin Kulldorff, (ii) the National Cancer Institute, and (iii) Farzad Mostashari of the New York City Department of Health and Mental Hygiene.

Latent class analysis

PROC LCA, developed by Lanza and colleagues, was used to conduct the latent class analysis. [112] In this analysis, the goal was to identify homogenous subgroups characterized by a combination of environmental emissions exposures using the dichotomized CMBG estimates of exposure. Zinc (Zn) and Selenium (Se) were also included, given recent findings that suggest they can moderate or reverse the impact of cadmium on the risk of HDP. [106] To determine the most appropriate class structure, following the guidance from Lanza, ten sets of models with random starting values consisting of two to five classes were run. The model was considered identified in sets where at least 80% of the models converged to the same solution. [112] Model fit was assessed using the Akaike information criterion (AIC), Bayesian information criterion (BIC), entropy, visual distinctiveness of each class, and class size. [112,113] Specifically,

models with the highest entropy, relatively low AIC and BIC, and class sizes greater than five records were prioritized.

Statistical Modeling

To assess the distribution of demographic characteristics, we summarized the overall sample with frequencies and percentages and provided row percentages for each demographic subgroup by exposure class. We used counts and column percentages to summarize the prevalence of eHDP among sociodemographic factors. A bivariate logistic regression was fit to examine the relationship between eHDP and covariates (sociodemographic factors and environmental exposure class). Variables were selected for inclusion into the final model, a multivariable logistic regression, if they were noted as being statistically associated with the prevalence of eHDP in the literature (maternal age, race, obesity, pre-existing diabetes, smoking), [6] were exposures of concern (As, Cd, Cr, Hg, and Pb), or were significantly associated with eHDP in the bivariate logistic regression (mother's ethnicity, education, ADI, Appalachian region, RUCC status, and stillbirth). Geocoding precision was included to adjust for geocoding misclassification. The final model consisted of latent class metal exposures, mother's age (years), mother's race, mother's ethnicity, mother's BMI, pre-existing diabetes, smoking through pregnancy, ADI, and the Appalachian region, Geocoding precision, RUCC and stillbirth To assess any potential biases that resulted from using the latent class assignment as a categorical variable, we conducted an additional multivariable logistic regression with a subset of the participants with a posterior probability (PPr) of less than 80%. The proportion of records in each class with a PPr of less than 80% are summarized in Appendix A.

SAS v 9.4 (SAS Institute, Cary, NC) was used for all non-spatial statistical analyses.

4.4 Results

4.4.1 Spatial Analysis

Figure 4.1 displays a series of maps that display the two individual level clusters of eHDP and choropleth maps of the median emissions for each chemical over the study period. Women living within the largest cluster, located within the Appalachian region, had a 63% greater risk of eHDP compared to women located outside the cluster (RR=1.63, 95% CI 1.38, 1.93). The second cluster, in Western Kentucky, was smaller, but women within this cluster had a 2-fold greater risk of eHDP compared to women outside of the cluster. (RR=2.21, 95% CI: 1.69, 2.90). Both clusters were significant for the entire duration of the study period (2007-2017). We also observed that, compared to other regions of the states, Louisville, Kentucky's largest city, had high median concentration estimates of As, Cr, Se, and Zn. The state's southwestern border (near Hopkinsville) had elevated median emissions of Zn, Se, and Cd. The majority of the Appalachian region in eastern Kentucky had low emission medians over the study period.

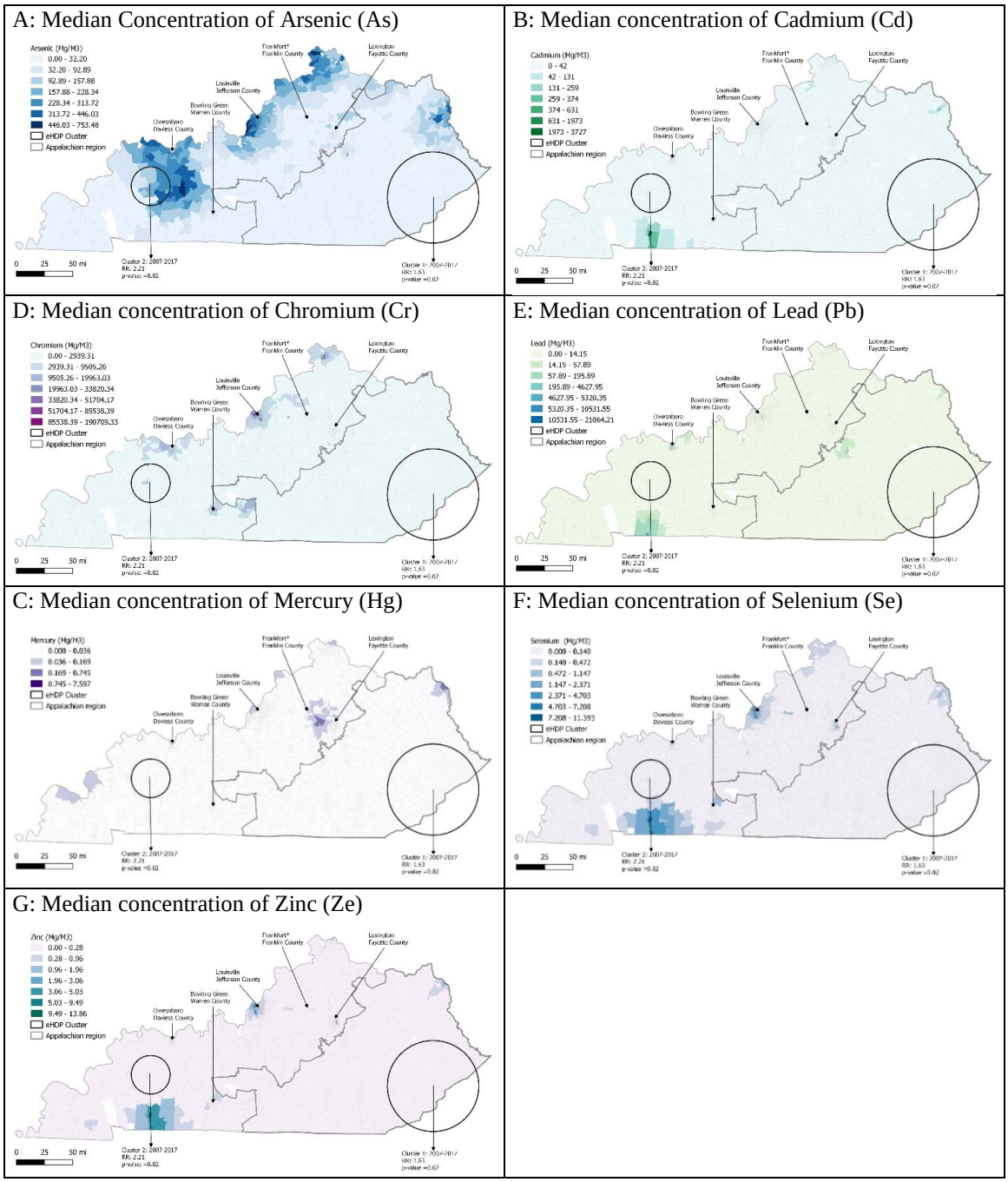


Figure 4.1 Median toxicity concentrations of emissions for each census micro-block group, Kentucky 2007-2017

4.4.2 LCA

Table 4.1 summarizes the model fit characteristics for LCAs with 2-4 classes. We chose the four-class model as the groups were distinctive, all classes had greater than five records, the AIC and the BIC were the lowest of the four models, and entropy was the

highest. Figure 4.2 visualizes the item response probabilities for each class. The first class (N=25,596) had a 74% probability of elevated Cd exposure, 62% probability of elevated As exposure, and a 74% probability of Pb exposure. Women in the second class (N=439,778) had a high probability of living in an area with both Se (91%) & Zn (86%) exposure. Those in the third exposure class (n=16,575) had a 99% probability of Pb exposure and a slightly elevated probability of Cr (51%). The final class, which constituted the majority of first-time mothers (n=119,687), had almost no metal exposure.

Table 4.1 Indicators of fit for latent class analysis with two through four classes of emission exposures

Number of Classes	AIC	BIC	Entropy
2	85797	85950.9	0.95
3	34648	34884.1	0.86
4	7388.41	7706.61	0.89

LCA: Latent Class Analysis, AIC: Akaike's Information Criteria.BIC: Bayesian Information Criteria;

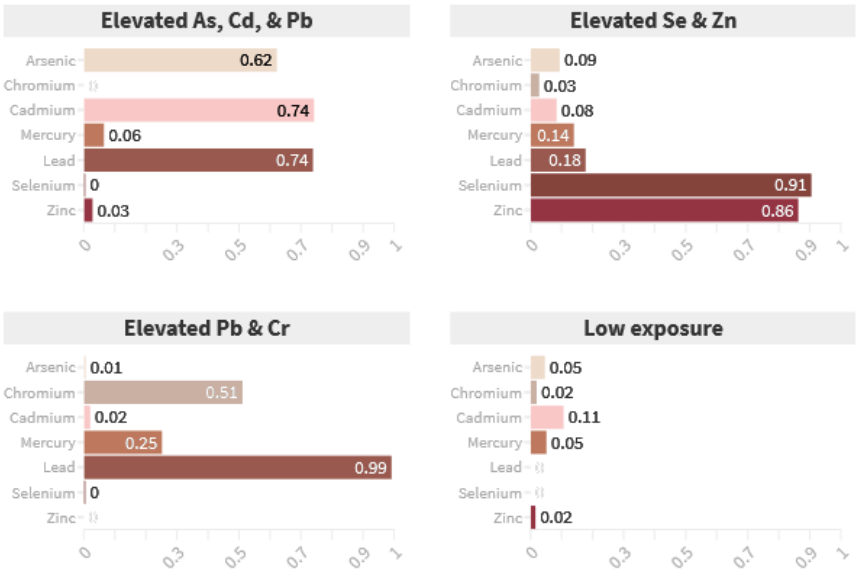


Figure 4.2 Class membership probabilities as a function of environmental chemical exposure

4.4.3 Statistical Modeling

Table 4.2 describes the demographic characteristics of the study population overall and by latent class. The sample predominantly was comprised of non-Hispanic (96%) White (85%) women between the ages of 21-24 (35%), with a BMI less than 25 (kg/m^2) (52%), and largely non-smoking (86%). The first class, which had high probabilities of As, Cd, & Pb exposure, contained 16% of Hispanic and 32% of Black women. Almost 9% of mothers that experienced eHDP were in this class. The second class, which had elevated Se & Zn exposures, included 24% of Hispanic women and 22% of White women. Approximately 22% of smokers fell in the second class, 19% of stillbirth records, and 22% of eHDP cases. The third class, defined by the high probability of elevated Pb & Cr exposure, had the smallest proportion of Black mothers (10%) of any metal exposure class and included 9% of eHDP records. The fourth class, the low exposure group, comprised 60 of those <20 years old and almost 60% of White mothers.

Table 4.2 Demographic characteristics summary by latent metal class and the total population of primiparous mothers 2008-2017

	Total	**Elevated As, Cd & Pb**	**Elevated Se & Zn**	**Elevated Pb & Cr**	**Low Exposure**
	205836 (100%)	25596 (12.4%)	43978 (21.4%)	16575 (8.1%)	119687 (58.2%)
	N (%)	***%***	***%***	***%***	***%***
Mother's age (years) †					
>35	10054 (4.7)	13.9	22.1	7.4	56.7
29-34	27686 (13.1)	13.8	22.6	7.5	56.2
25-28	55338 (26.1)	12.5	23.1	8.2	56.3
21-24	73473 (34.7)	11.9	20.6	8.4	59.0
<20	45500 (21.5)	13.0	19.8	7.7	59.5
Mothers race†					
Black	17512 (8.3)	31.4	14.8	10.1	43.7
Other	14222 (6.7)	18.8	21.0	8.2	51.9
White	180317 (85.0)	10.3	22.1	7.8	59.8
Mother's ethnicity†					
Non-Hispanic	203521 (96.0)	12.5	21.3	8.0	58.2
Hispanic	8530 (4.0)	16.3	24.2	8.0	51.4
Mother's education†					
College Degree	69526 (33.1)	12.6	23.5	7.9	56.0
Some College	49760 (23.7)	13.2	21.5	8.8	56.5
High School	58018 (27.6)	10.8	20.5	7.9	60.8
Less than HS	32824 (15.6)	14.8	18.5	7.6	59.1
Mother married†					
Yes	107478 (50.7)	10.7	22.5	8.6	58.3
No or not stated	104573 (49.3)	14.7	20.3	7.5	57.5

Mother's BMI (kg/m^2)†					
Obese (≥30)	49608 (24.1)	11.6	20.9	7.4	60.1
Overweight (25-30)	50122 (24.4)	12.4	21.4	7.9	58.4
Underweight/Normal (<25)	106106 (51.6)	12.9	21.6	8.5	57.1
Pre-existing diabetes †					
Yes	1608 (0.8)	10.8	21.3	7.2	60.8
No or not stated	210443 (99.2)	12.6	21.4	8.1	57.9
Smoking throughout pregnancy†					
No or not stated	182021 (85.8)	13.0	21.7	8.1	57.2
Yes	30030 (14.2)	10.5	19.6	7.8	62.2
ADI†					
Impoverished*	33473 (15.8)	12.5	8.9	3.7	74.9
No/lower impoverishment	178578 (84.2)	12.7	23.8	8.9	54.8
Appalachian†					
Appalachian	53904 (25.4)	0.3	8.1	10.1	81.4
Not Appalachian	158147 (74.6)	16.8	25.9	7.3	49.9
Geocoding precision†					
Address point/Street segment	189367 (89.3)	13.5	22.3	8.3	55.9
Imprecise**	22684 (10.7)	5.7	13.8	5.9	74.7
RUCC Status†					
Rural	17261 (8.1)	0.0	5.9	2.9	91.2
Non-metro	66531 (31.4)	0.5	15.8	4.4	79.3
Urban	128259 (60.5)	20.6	26.4	10.6	42.4

Stillbirth†					
Yes	1247 (0.6)	14.4	18.6	7.3	59.7
No or not stated	210804 (99.4)	12.6	21.4	8.1	57.9

eHDP: an early-onset hypertensive disorder of pregnancy, where hypertensive symptoms present before 35 weeks; BMI: Body Mass Index; RUCC rural-urban continuum codes, Metal Abbreviations: As: Arsenic, Cd: Cadmium, Cr: Chromium, Hg: Mercury, Pb: Lead, Se: Selenium, Zn: Zinc

*Impoverished Upper quintile of economic deprivation (9-10) **Imprecise address: Midpoint of street/ City/Zip/ No Geocode,

† Chi-square test statistics <0.05,

Table 4.3 Bivariate and multivariable associations between demographic characteristics, environmental exposures, and class membership

	eHDP N (%)	PR (95% CI)	aPR (95% CI)	Sensitivity aPR (95% CI)
Latent class metal exposure				
Elevated As, Cd & Pb	166 (8.6)	0.64 (0.56, 0.76)	0.72 (0.60, 0.86)	0.82 (0.20, 3.46)
Elevated Se & Zn	418 (21.6)	0.96 (0.86, 1.08)	1.08 (0.96, 1.22)	0.82 (0.50, 1.32)
Elevated Pb & Cr	180 (9.3)	1.12 (0.94, 1.30)	1.22 (1.04, 1.44)	1.58 (1.02, 2.46)
Low Exposure	1168 (60.5)	Reference	Reference	Reference
Mother's age (years)				
≥35	142 (7.3)	2.06 (1.70, 2.52)	1.96 (1.54, 2.50)	2.68 (1.20, 5.98)
29-34	289 (15.0)	1.52 (1.30, 1.78)	1.54 (1.26, 1.88)	1.68 (0.82, 3.46)
25-28	541 (28.0)	1.42 (1.24, 1.64)	1.38 (1.16, 1.66)	1.26 (0.68, 2.36)
21-24	647 (33.5)	1.28 (1.12, 1.46)	1.12 (0.96, 1.30)	1.38 (0.82, 2.30)
≤20	313 (16.2)	Reference	Reference	Reference
Mother's race				
Black	237 (12.3)	1.52 (1.32, 1.74)	1.60 (1.36, 1.88)	2.08 (1.28, 3.34)
Other	76 (3.9)	0.60 (0.48, 0.74)	0.72 (0.56, 0.94)	0.18 (0.02, 1.32)
White	1619 (83.8)	Reference	Reference	Reference
Mother's ethnicity				
Non-Hispanic	1879 (97.3)	1.50 (1.14, 1.96)	1.04 (0.76, 1.44)	2.42 (0.32, 18.84)
Hispanic	53 (2.7)	Reference	Reference	Reference
Mother's education				
College Degree	612 (32.0)	1.40 (1.20, 1.64)	1.12 (0.92, 1.38)	0.86 (0.40, 1.78)
Some College	540 (28.3)	1.74 (1.48, 2.04)	1.38 (1.14, 1.66)	1.50 (0.78, 2.88)
High School	553 (28.9)	1.52 (1.30, 1.78)	1.28 (1.08, 1.52)	1.34 (0.72, 2.48)
Less than HS*	206 (10.8)	Reference	Reference	Reference

Mother's BMI (kg/m^2)				
Obese (≥30)	766 (42.4)	2.80 (2.50, 3.12)	2.44 (2.18, 2.72)	3.06 (2.06, 4.54)
Overweight (25-30)	448 (24.8)	1.60 (1.42, 1.82)	1.52 (1.34, 1.72)	2.04 (1.34, 3.14)
Underweight/Normal (<25)	592 (32.8)	Reference	Reference	Reference
Pre-existing diabetes				
Yes	87 (4.5)	6.46 (5.18, 8.06)	4.62 (3.64, 5.84)	3.64 (1.40, 9.50)
No or not stated	1845 (95.5)	Reference	Reference	Reference
Smoking throughout pregnancy				
No or not stated	1700 (88.0)	1.22 (1.06, 1.38)	1.30 (1.10, 1.50)	1.74 (0.98, 3.10)
Yes	232 (12.0)	Reference	Reference	Reference
ADI				
Impoverished area (9-10)	353 (18.3)	1.20 (1.06, 1.34)	1.08 (0.94, 1.22)	1.46 (0.82, 2.6)
No/lower impoverishment (1-8)	1579 (81.7)	Reference	Reference	Reference
Appalachian				
Appalachian	599 (31.0)	1.32 (1.20, 1.46)	1.20 (1.04, 1.36)	0.66 (0.38, 1.16)
Not Appalachian	1333 (69.0)	Reference	Reference	Reference
Geocoding precision				
Address point/Street segment	1701 (88.0)	0.88 (0.76, 1.02)	0.98 (0.84, 1.14)	1.18 (0.64, 2.24)
Imprecise**	231 (12.0)	Reference	Reference	
RUCC Status				
Rural	169 (8.7)	1.20 (1.02, 1.42)	1.06 (0.86, 1.28)	2.60 (0.78, 8.70)
Non-Metro	720 (37.3)	1.34 (1.22, 1.46)	1.20 (1.04, 1.36)	1.60 (1.06, 2.44)
Urban	1043 (54.0)	Reference	Reference	Reference
Stillbirth				
Yes	56 (2.9)	5.24 (4, 6.88)	4.30 (3.20, 5.78)	7.42 (3.06, 18.02)
No or not stated	1876 (97.1)	Reference	Reference	Reference

BMI: Body Mass Index, ADI: Area Deprivation Index, RUCC: Rural-Urban Continuum Codes,
*HS: High School
**Imprecise address: Midpoint of street/ City/Zip/ No Geocode,

Table 4.3 summarizes the proportion of eHDP cases by covariate and displays the results of the bivariate, multivariate, and sensitivity analyses with posterior probabilities (PPr) of less than 80%. Please see the appendix for counts, percentages, and average PPr of each class for the overall sample and those with a PPr of less than 80%.

Approximately 22% of women who experienced eHDP were in the second latent class, with high probabilities of exposures to elevated Se & Zn (22%). Women who experienced eHDP were predominantly White (84%) non-Hispanic (97%) between 21-24 years old (34%). Over 42% were obese, 12% reported smoking throughout their pregnancy, and 52% reported being married.

In the bivariate assessment, women in the first class with a high probability of concurrent exposure to As, Cd, & Pb, had a 36% reduction in the prevalence of eHDP (PR=0.64, 95% CI: 0.56, 0.76) compared to women with low metal exposure. In contrast, women in the third class, with concurrent elevated Pb & Cr exposure, had a 22% increase in the prevalence of eHDP (PR=1.12, 95% CI: 0.94, 1.30). Age was also associated with the prevalence of eHDP, with women $\geq$ 35 years old having twice the prevalence of eHDP compared to those $\leq$ 20 years old (PR=2.06, 95% CI: 1.70, 2.52). Mothers with an obese BMI had an almost three times higher prevalence of eHDP than those underweight/normal BMI (PR=2.80, 95% CI: 2.50,3.12). Black women also experienced a 52% higher prevalence of eHDP (PR=1.52, 95% CI: 1.32, 1.74) than White women.

Following covariate adjustment, those in the first latent class with elevated As, Cd, & Pb had a 38% reduction in eHDP compared to those with low exposure (aPR=0.72, 95% CI: 0.60 0.86), and those in the third class with an elevated Pb & Cr experience had 22% increase prevalence of eHDP (aPR=1.22, 95% CI: 1.04, 1.44). The prevalence of eHDP in those $\geq$ 35 years old remained high after adjustment, compared to those $\leq$ 20 years old (aPR=1.96, 95% CI: 1.54, 2.50). After adjustment, the prevalence of eHDP among those that were obese declined slightly (aPR=2.44, 95% CI: 2.18, 2.72) but still two times higher among women with an obese BMI. Prevalence estimates were similar among age, race, smoking throughout pregnancy, stillbirth, and among geographic covariates such as Appalachian status, geocoding precision, and RUCC status. For those who had pre-existing diabetes, there was a marked decline in the

prevalence of eHDP after adjustment (PR=6.46, 95% CI: 5.18, 8.06 vs. aPR=4.62, 95% CI: 3.64, 5.84).

Most estimates in the sensitivity analysis were within the bounds or did not deviate substantially from the adjusted analysis. However, the mother's race saw the prevalence ratios become more extreme (sensitivity aPR=2.1, 95% CI 1.28, 3.34), as did the aPR for ADI (aPR =1.46 95% CI: 0.82, 2.60). In Appalachia, the prevalence of eHDP reversed from the adjusted multivariable logistic regression (sensitivity aPR=0.66, 95% CI: 0.38, 1.16).

4.5 Discussion

This study sought to identify geospatial trends and exposure patterns of environmental air emissions of As, Cd, Cr, Hg, and Pb in Kentucky and assess the relationship between these exposures and eHDP, adjusting for demographic risk factors for eHDP. Our findings contribute to emerging literature on the association of metal exposures with HDP, specifically, eHDP. Using individual-level birth records, we discovered two statistically significant clusters of eHDP, one in Western Kentucky and a second larger cluster in the Appalachian region. Employing an LCA, we identified four subgroups of metal exposures and further detected that women in the latent class with elevated exposure to Pb & Cr had a statistically significant increase in the prevalence of eHDP after covariate adjustment. [114] We also found that individual factors such as Black race, maternal age $\geq$ 34 years, obesity, and smoking throughout the pregnancy were associated with an increased prevalence of eHDP. Non-metro women or those who lived within the Appalachian region also had a higher prevalence of eHDP. Demographic covariates that strongly influence the prevalence of eHDP support most of the existing literature, with smoking being a notable exception. [70,71]

Although elevated emissions were largely in more urban areas and along the border, the eHDP cluster in Western Kentucky overlapped with census micro-block groups (CMBG) with elevated median estimates of As. However, the second cluster, located in the eastern, Appalachian region of the state, had very little visual overlap with the examined metal emissions. These clusters may reflect the geospatial patterns of eHDP

risk factors, such as obesity and pre-existing diabetes, which are particularly elevated in the Appalachian region.[112] However, further study, particularly assessing the specific environmental characteristics of the cluster areas, is needed.

While there have been limited studies assessing the impact of environmental metal exposures, specifically those identified in our study – Pb and Cr – on eHDP, individual-level assessments suggest a relationship between exposure and eHDP. Porporat and colleagues concluded in a recent literature review that leads exposure is among the most important risk factors for pre-eclampsia yet identified. [115] In a cohort study, Yazbeck and colleagues found a statistically significant relationship between blood levels of Pb at 24-28 weeks gestation was significantly higher in women with HDP than their non-HDP counterparts. Chromium has had less clear results. Researchers noted that pregnancies in areas with drinking water contaminated with metals such as Cr, Pb, and As, that women had elevated odds of HDP. [32] A case-control study of South African women found cases had significantly elevated levels of chromium in pubic hair compared to controls. [116]

To assess environmental metal exposures, we employed toxicity concentration estimates from RSEI, a promising publically available population-level dataset that, among other things, estimates the environmental volume of chemicals of concern at small spatial scales (810 m x 810 m grids, CMBG, census tracts, and ZIP codes). Although RSEI data has had a limited assessment in health research, preliminary work has found that children living in areas with RSEI estimated elevated aerosolized concentrations of Hg and Cd and reported elevated blood metal levels. [117] In a study assessing geospatial patterns and risk factors for preterm birth, Ogneva-Himmelberger and colleagues reported an association between pre-term births and RSEI estimated hazard scores. [118]

However, the use of RSEI data is not without limitations. These data are based on Toxic Release Inventory (TRI) reports, which are self-disclosed yearly estimates of fugitive and stack emissions by only a subset of facilities located in the US. Although site-specific characteristics are incorporated if available, generalizations are often made to similar facilities' characteristics if not. The RSEI model also does not incorporate non-TRI sources of contamination and does not examine, integrate, or estimate decay

products of emissions, which may not have equivalent health risks as the parent product. However, RSEI data offer estimates of environmental concentrations of almost 700 specific contaminants of concern in both air and water (this study focused exclusively on air), adjusted for characteristics of the contaminant at a spatially confined resolution. Estimates are updated based on the most current data (facilities can correct reported emissions information for up to 3 years) and the most up-to-date methodologies. Changes in reporting standards have also been relatively few since the program's inception in 1988, making the data appropriate for longitudinal analyses.

The use of RSEI data, which employs consistent methodologies to estimate the local environmental burden of chemicals of concern, also allowed for the exploration of novel methodologies to assess concurrent environmental exposures to chemicals of concern. Traditionally, to assess the impact of environmental emissions exposure on poor pregnancy outcomes, studies employ a generalized linear model to individually assess the impact of exposures, adjusting for other important risk factors, on the outcome of interest, with limited (if any) assessment of interactions. Although helpful, these statistical models may not adequately address the impact of interactions among various chemicals, which rarely occur in isolation. To explore concurrent environmental exposures, we employed a latent class model to identify patterns of exposure. The latent class analysis (LCA) creates homogenous and mutually exclusive subgroups using the similarities of response patterns among records. [119] This person-centered approach allows for evaluating complex interactions without sacrificing statistical power, as multiple exposures are combined based on the probability of concurrent exposure. However, these classifications may not accurately reflect subgroups within the population, and records may be inappropriately assigned to exposure classifications if they do not have a clear class assignment. To assess potential bias in latent classes, we conducted a sensitivity analysis for those with low posterior probabilities ($\leq$80%). We found that, although the point estimates for measures of association were more extreme, the majority of prevalence ratios were in the same direction and had overlapping confidence intervals. However, a few covariates reversed direction (elevated Se & Zn exposure, college education, Appalachian status, and address point/street segment geocoding precision). Overall, this suggests although there was some bias, the impact appears to be limited.

Although this study has notable strengths, such as assessing all of the primiparous births in Kentucky over ten years (2008-2017), there are significant limitations. Birth certificates may be subject to recall and interviewer bias, particularly with health information such as smoking and pre-pregnancy weight. Additionally, although the geocoding precision of Kentucky birth certificates has been found to be adequate, previous analyses have found that White women in urban areas and Black women in rural areas have an increased prevalence of geocoding imprecisely (see Chapter 2 for further details). Moreover, we could not determine the mother's length of time at the residential address provided on the birth certificate. [93]

Also of concern is the impact of geocoding imprecision on exposure classification and spatial cluster detection. In this study, we determined the CMBG based on the geographic coordinates of the geocoded maternal addresses. The CMBG was then used to assign exposure status and assess the geospatial clustering of eHDP. Overall, 90% of records used in this study were geocoded precisely; within the non-Appalachian region, the proportion of precise addresses was almost 95%, which supports our belief that the clusters detected are accurate, and there is limited exposure misclassification.

In contrast, however, almost 24% of addresses within the Appalachian region geocoded imprecisely; that is to either the midpoint of the street, ZIP code, or city. As the emission patterns of environmental metals were homogenous across the Appalachian region, we do not suspect substantial misclassification of exposure. However, we encourage a cautious interpretation of the spatial cluster identified within the Appalachian region. Although we believe that the Appalachian region has a high burden of eHDP, the cluster identified in this study may not be precise. Further study in the cluster area and neighboring communities is needed.

In addition to improving the GIS resources within the Appalachian region, future assessments focusing on evaluating a broader range of exposure sources such as drinking water, residential air quality, and occupational exposures would allow for a fuller picture of metal exposures and lead to a better evaluation of health outcomes. Additionally, further exploration of the interaction among toxic exposures would help expand some of

the results observed in this study, such as the protective effect of As, Cd, & Pb, but the increased risk for those with a high probability of Pb & Cr exposure.

4.5.1 Conclusions

This study adds to the limited literature examining the risk of HDP by focusing explicitly on early-onset HDP and employing a latent class methodology to assess multiple environmental exposure patterns. Further, we identified metal exposures, specifically Pb & Cr exposures, as contributing to the prevalence of eHDP. Findings from this study suggest that efforts to mitigate metal exposures among women of childbearing age are highly warranted.

CHAPTER 5. DISCUSSION

Hypertensive disorders of pregnancy (HDP) are increasing in the United States, with noted elevation in Kentucky. [7,9] These disorders are among the leading causes of maternal and infant mortality in the morbidity in the US, and associated healthcare costs exceed over $6 billion in the first year after birth. [6,120] Early-onset hypertensive disease (eHDP), a subset of HDP, is particularly concerning, as it is associated with the highest risks of maternal death and long-term chronic conditions. [11,54,60] Further, clinical research has indicated early-onset disease may not share the same risk factors of late-onset disease, although some risk factors such as obesity, pre-existing diabetes, and advanced maternal age; appear to be consistent across classification, even if the strength of the association varies among disease classifications. [56,121] There has been limited exploration of environmental risk factors; however, early results indicate associations with environmental exposures and HDP. [6,14-21] With the increased prevalence of risk factors such as obesity and advancing maternal age in primiparous mothers, it is crucial that we characterize the disease burden of eHDP, one of the most severe forms of HDP, and assess environmental risk factors associated with disease to improve maternal and infant health outcomes.

To better characterize eHDP in Kentucky and explore potential geospatial trends, we used Kentucky birth records (2008-2017) and merged geocoded maternal addresses with data from the environmental protection agency's (EPA) Risk Screening Environmental Indicators (RSEI) model (2007-2017) to assess environmental exposures. However, before merging these datasets, the geocoding precision of the address records needed to be characterized to discern and articulate limitations in future studies. This process and the geocoding precision of Kentucky birth records are discussed in chapter three, "Geocoding precision of birth records from 2008 to 2017 in Kentucky, USA." In the second chapter, "County prevalence and geospatial trends of early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017," we assess county-level geospatial and temporal trends and identify covariates associated with eHDP. The fourth chapter, "A cross-sectional examination of the association of environmental metal exposure with early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017", explores

patterns of exposure to aerosolized metals through industrial activity (specifically, arsenic, cadmium, chromium, lead, and mercury) and the association with eHDP, adjusting for geocoding precision.

The first study of this dissertation, "Geocoding precision of Kentucky birth records, 2008-2017," assessed geospatial patterns of geocoding precision and identified covariates associated with poor geocoding precision to increase the transparency of our analytic procedures and to understand potential bias in future studies, as the later projects would use geocoded coordinates to assess county-level prevalences and assign exposure status. Previous studies have found that geocoding precision is impacted by both population density, as urban areas tend to geocode with more precision than rural areas, and demographic characteristics, such as race – factors often associated with disease status. [37,51] In a meta-analysis of electronic health records using GIS methods, researchers noted that less than one-third of journal articles described their geocoding process. [40,46,122] This lack of disclosure prevents other researchers from assessing the potential for bias in any study that uses geocoded address data. We found that Kentucky vital statistics data addresses geocoded precisely in metropolitan areas but lower in non-metro and rural areas. However, there was an improvement over the study period. We also identified disparities in geocoding precision among each of the rural-urban continuum codes (RUCCs). Specifically, in metropolitan areas, the addresses of White, non-Hispanic mothers had the highest odds of poorly geocoded addresses; however, in rural areas, Black or Hispanic mothers had a higher odds of imprecise geocoding. Future work exploring geospatial trends would benefit from assessing patterns of geocoded precision, particularly within racial groups and population density metrics, such as RUCC designations, as both urban-metro status and race have been associated with the odds of geocoding imprecisely.

The second study, "County prevalence and geospatial trends of early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017," explores temporal and geospatial patterns of eHDP and identifies demographic factors associated with an increased prevalence of eHDP. There has been limited assessment of disease prevalence based on the gestational age of symptom onset, as most studies focus on diagnoses of any HDP or general pre-eclampsia. However, the rates of HDP, which include eHDP, are

concerning. In a study on trends of non-specific pre-eclampsia across the U.S. from 1996-2004, Wallis and colleagues reported that the rate of P.E. in the Southern portion of the U.S. was approximately 34.1 per 1000 live births, a 12 higher prevalence compared to the Northeast region of the U.S. (RR=1.12, 95% CI: 1.03, 1.21). [8] In a one-year cross-sectional study of 2017 birth data, Butwick and colleagues reported that Kentucky had approximately 70 cases of HDP per 1000 live births – the 8th highest prevalence in the US. [7] In the second study, which focused on early-onset disease manifestation in Kentucky, found that the prevalence of eHDP was, on average, 9.2 cases per 1,000 births across all counties during the study period and that there was an almost 3% rate increase during the study period (2008-2017). Although there are no nationwide estimates of eHDP prevalence, a study using birth records from Washington reported 3.8 eHDP per 1,000 births – slightly less than half of Kentucky rates. [57] These findings may reflect the general elevation of obesity and pre-existing diabetes in Kentucky relative to Washington; however, further study is needed, as other factors, such as environmental regulations and predominant industry, vary between these states. [79,80]

Variations in HDP rates in the U.S. have been attributed to the geographic variation in health behaviors and pre-existing conditions such as obesity and smoking; however, there has been very little assessment of non-clinical explanations for rate changes. [8,9] We found that elevated county-level proportions of obesity, current cigarette smoking, and marriage were associated with an increased prevalence of eHDP. These findings may be due to the limitations inherent within ecological studies, notably the ecological fallacy; however, poor reproductive health literacy among young adults, especially in high-poverty regions, may also contribute to these findings.

In the fourth chapter, “Cross-sectional analysis of Risk Screening Environmental Indicators toxicity concentration estimates and early-onset hypertensive disorders of pregnancy in Kentucky, 2008-2017,” we focused on the incidence of eHDP, adjusting for, among other covariates, geocoding precision. In this study, we sought to identify geospatial trends and exposure patterns of environmental air emissions of arsenic, cadmium, chromium, lead, and mercury in Kentucky and describe the prevalence of eHDP according to levels of toxic metal exposure. Using individual-level data, we discovered two statistically significant clusters of eHDP, one in Western Kentucky and a

second larger cluster in the Appalachian Region. Employing a latent class analysis (LCA), we identified four subgroups of metal exposures. Women in the latent class with elevated probability of exposure to lead and chromium had a statistically significant increase in the prevalence of eHDP after covariate adjustment. [114] We also observed that the prevalence of eHDP in Kentucky, 9.1 per 1000 births, was higher than reported in a similar study conducted in Washington (3.8 per 1000 live births). There was no statistically significant association between geocoding precision and eHDP prevalence, suggesting that although there is bias in geocoding precision, it is not strongly associated with eHDP (see Chapter 2). These findings may reflect the general elevation of obesity and pre-existing diabetes in Kentucky relative to Washington. [79,80] Differences in the distribution of risk factors within the state may also explain the location of these clusters in the Western and Appalachian regions of the state, as rural areas have a higher prevalence of obesity and pre-existing diabetes. [103]

This study contributes to an emerging literature on the association of environmental exposures with HDP, specifically, early-onset HDP. The use of RSEI data, which provides estimates of air and water chemical toxicity concentrations for each 810 m x 810 m grid across the United States, can provide precise estimates of industrially emitted exposures of concern, with the flexibility of scaling to larger spatial areas, such as the census micro-block groups used in this study. The use of the LCA allowed for the assessment of patterns of metal exposure across the study region. This person-centered approach creates groups based on the similarities of response patterns, which allows for the evaluation of complex interactions without sacrificing statistical power. [119]

5.1 Strengths and limitations

Birth and RSEI data, although useful, have several noted limitations. First, birth data has been shown to underreport maternal and infant complications, and we could not verify health or address data. [123] Further, a recent study found substantial variation in collection procedures in Kentucky hospitals. Also of note is the inability to explore the association of chronic hypertension with eHDP, as on the birth certificates, chronic hypertension and HDP are treated as mutually exclusive. Additionally, this dissertation focused on those that gave birth before 34 weeks, which invariable missed cases and

could have biased our estimates, as these individuals had the most severe disease. The environmental exposure data, RSEI, likewise has notable shortcomings. Foremost is self-report bias, as RSEI obtains information on volumes emitted from reports submitted to the Toxic Release Inventory (TRI) program. These may not be verifiable, as TRI guidelines state that estimates are sufficient. [109] RSEI may incorporate additional estimates within the modeling protocol if specific facility information (such as stack heights) is unavailable. Estimates of environmental chemical burden also do not account for product decay or sources of exposure outside of TRI reporting.

Strengths of this study include using the majority of births to Kentucky residents to estimate prevalence rates accurately. In addition, the high precision of geocoding within the population indicates that for most areas and people, we are able to locate their residence precisely. Therefore exposure estimates, which were on a microblock group level – are more likely to reflect the environmental conditions of those residential areas. Finally, although there are noted concerns of RSEI data, their methodologies are consistent throughout the entire study period.

5.2 Future work

Future work is needed given the prevalence of risk factors associated with eHDP and the association with environmental exposures and eHDP risk. First and foremost, a more accurate assessment of HDP in Kentucky is needed. This study focused on those who gave birth before 34 weeks, as the birth data could not ascertain the gestational age of symptom onset. Future research should accurately identify disease and gestational age of onset to identify risk factors and more accurately describe cases.

This research also identified clusters of disease with an elevated prevalence compared to the rest of the state. A more thorough health and family history assessment and accurate characterization of local environmental and occupational exposures of primiparous women and their partners living in these areas could further elucidate causal factors associated with eHDP.

APPENDIX A

To further assess the LCA results presented in chapter 3, we also assessed the posterior probabilities or the probability that a record would be assigned to a given class. To ensure that the classes were sufficiently distinctive, we calculated each class's average posterior probability (PPr). The average lowest posterior probability, which was the low exposure group, was 86%, with an average of 85% indicating adequate class assignments.

There were 19,087 (9%) records with a posterior probability of less than 80%; the proportion by latent class is displayed in Table 4. Most of the records were assigned a posterior probability >80% for the latent class assignment, although those in the low exposure class had the lowest proportion (70.1%) of records assigned with confidence.

Table A 1 Summary of Latent metal class and average posterior probability for class assignment

		Total	**<80% Posterior Probability**	
Class description	**N(%)**	**X (SD)**	**N (%)**	**X (SD)**
1. Elevated As, Cd, & Pb	26773 (98.5)	96.4 (5.9)	411 (1.5)	70.2 (3.9)
2. Elevated Se & Zn	45391 (91.0)	95.1 (13.1)	4514 (9.1)	57.1 (5.1)
3. Elevated Pb & Cr	122829 (94.7)	95.9 (6.5)	6896 (5.3)	78.1 (0.4)
4. Low Exposure	17058 (70.1)	86.4 (13.9)	7266 (29.9)	70.5 (1.1)

APPENDIX B

This table presents the geocoding precision of the records before excluding records that geocoded outside of the state. (For Chapter 4)

		Total N (%)	Non-Appalachian N (%)	Appalachian N (%)
High Coordinate Precision	 **Address Point**: An exact match of the address has been found	127,165 (60.0)	105,681 (66.8)	21,484 (39.9)
	 **Street Segment:** The address has been narrowed down to a short segment of a street	63,960 (30.2)	44,234 (28.0)	19,726 (36.6)
Low Coordinate Precision	 **Street:** Street has been identified, but house number may not be in the range of the street.	3,879 (1.8)	1,608 (1.0)	2,271 (4.2)
	 **ZIP/City**	17047 (8.0)	6,624 (4.2)	10,423 (19.3)
	 **Total records**	N = 212,527	N = 158,147	N = 53,904

Figure B.1: Description and summary of geocoding precision for Kentucky addresses, 2008-2017

References

1. Pregnancy Mortality Surveillance System | Maternal and Infant Health | CDC. 2020; https://www.cdc.gov/reproductivehealth/maternal-mortality/pregnancy-mortality-surveillance-system.htm. Accessed 9/11/2021, 2021.
2. ACOG Practice Bulletin No. 203: Chronic Hypertension in Pregnancy. *Obstet Gynecol.* 2019;133(1):e26-e50.
3. Ying W, Catov JM, Ouyang P. Hypertensive Disorders of Pregnancy and Future Maternal Cardiovascular Risk. *Journal of the American Heart Association.* 2018;7(17):e009382.
4. Gestational Hypertension and Preeclampsia: ACOG Practice Bulletin, Number 222. *Obstet Gynecol.* 2020;135(6):e237-e260.
5. Task Force on Hypertension in Pregnancy. *Hypertension in Pregnancy.* Washington, DC2013.
6. ACOG Practice Bulletin No. 202: Gestational Hypertension and Preeclampsia. *Obstet Gynecol.* 2019;133(1):e1-e25.
7. Butwick AJ, Druzin ML, Shaw GM, Guo N. Evaluation of US State–Level Variation in Hypertensive Disorders of Pregnancy. *JAMA Network Open.* 2020;3(10):e2018741-e2018741.
8. Wallis AB, Saftlas AF, Hsia J, Atrash HK. Secular trends in the rates of preeclampsia, eclampsia, and gestational hypertension, United States, 1987-2004. *American journal of hypertension.* 2008;21(5):521-526.
9. Ananth CV, Basso O. Impact of Pregnancy-Induced Hypertension on Stillbirth and Neonatal Mortality in First and Higher Order Births: A Population-Based Study. *Epidemiology (Cambridge, Mass).* 2010;21(1):118-123.
10. Sibai B. HELLP Syndrome. 2018; https://www.uptodate.com/contents/hellp-syndrome. Accessed March 24, 2018, 2018.
11. Lisonkova S, Sabr Y, Mayer C, Young C, Skoll A, Joseph KS. Maternal morbidity associated with early-onset and late-onset preeclampsia. *Obstetrics and gynecology.* 2014;124(4):771-781.
12. Iacobelli S, Bonsante F, Robillard PY. Comparison of risk factors and perinatal outcomes in early onset and late onset preeclampsia: A cohort based study in Reunion Island. *J Reprod Immunol.* 2017;123:12-16.
13. Catov JM, Ness RB, Kip KE, Olsen J. Risk of early or severe pre-eclampsia related to pre-existing conditions. *Int J Epidemiol.* 2007;36(2):412-419.
14. Liu T, Zhang M, Guallar E, et al. Trace Minerals, Heavy Metals, and Preeclampsia: Findings from the Boston Birth Cohort. *Journal of the American Heart Association.* 2019;8(16):e012436.
15. Rosen EM, Munoz MI, McElrath T, Cantonwine DE, Ferguson KK. Environmental contaminants and preeclampsia: a systematic literature review. *Journal of toxicology and environmental health Part B, Critical reviews.* 2018;21(5):291-319.
16. Athukorala C, Rumbold AR, Willson KJ, Crowther CA. The risk of adverse pregnancy outcomes in women who are overweight or obese. *BMC Pregnancy and Childbirth.* 2010;10(1):56.
17. Cleary-Goldman J, Malone FD, Vidaver J, et al. Impact of Maternal Age on Obstetric Outcome. 2005;105(5):983-990.
18. Ananth CV, Keyes KM, Wapner RJ. Pre-eclampsia rates in the United States, 1980-2010: age-period-cohort analysis. *BMJ (Clinical research ed).* 2013;347:f6564.

19. Barton JR, Sibai AJ, Istwan NB, Rhea DJ, Desch CN, Sibai BM. Spontaneously conceived pregnancy after 40: influence of age and obesity on outcome. *American journal of perinatology.* 2014;31(9):795-798.
20. Tanaka M, Jaamaa G, Kaiser M, et al. Racial Disparity in Hypertensive Disorders of Pregnancy in New York State: A 10-Year Longitudinal Population-Based Study. *American Journal of Public Health.* 2007;97(1):163-170.
21. Ghosh G, Grewal J, Männistö T, et al. Racial/ethnic differences in pregnancy-related hypertensive disease in nulliparous women. 2014;24(3):283.
22. Elongi Moyene J-P, Scheers H, Tandu-Umba B, et al. Preeclampsia and toxic metals: a case-control study in Kinshasa, DR Congo. *Environ Health.* 2016;15:48-48.
23. Laine JE, Bailey KA, Rubio-Andrade M, et al. Maternal arsenic exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) pregnancy cohort in Mexico. *Environmental health perspectives.* 2015;123(2):186-192.
24. Sandoval-Carrillo A, Mendez-Hernandez EM, Antuna-Salcido EI, et al. Arsenic exposure and risk of preeclampsia in a Mexican mestizo population. *BMC Pregnancy Childbirth.* 2016;16(1):153.
25. Pan S, Lin L, Zeng F, et al. Effects of lead, cadmium, arsenic, and mercury co-exposure on children's intelligence quotient in an industrialized area of southern China. *Environmental pollution (Barking, Essex : 1987).* 2018;235:47-54.
26. Lamb MR, Janevic T, Liu X, Cooper T, Kline J, Factor-Litvak P. Environmental lead exposure, maternal thyroid function, and childhood growth. *Environmental research.* 2008;106(2):195-202.
27. Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. Heavy metal toxicity and the environment. *Exp Suppl.* 2012;101:133-164.
28. Claus Henn B, Ettinger AS, Hopkins MR, et al. Prenatal Arsenic Exposure and Birth Outcomes among a Population Residing near a Mining-Related Superfund Site. *Environmental health perspectives.* 2016;124(8):1308-1315.
29. Hu H, Shine J, Wright RO. The Challenge Posed to Children's Health by Mixtures of Toxic Waste: The Tar Creek Superfund Site as a Case-Study. *Pediatric Clinics of North America.* 2007;54(1):155-175.
30. Abhyankar LN, Jones MR, Guallar E, Navas-Acien A. Arsenic exposure and hypertension: a systematic review. *Environmental health perspectives.* 2012;120(4):494-500.
31. Pedersen M, Halldorsson TI, Olsen SF, et al. Impact of Road Traffic Pollution on Pre-eclampsia and Pregnancy-induced Hypertensive Disorders. *Epidemiology (Cambridge, Mass).* 2017;28(1):99-106.
32. Padula AM, Ma C, Huang H, Morello-Frosch R, Woodruff TJ, Carmichael SL. Drinking water contaminants in California and hypertensive disorders in pregnancy. *Environ Epidemiol.* 2021;5(2):e149-e149.
33. Agency for Toxic Substances and Disease Registry (ASTDR). In:2021.
34. Ash M, Fetter TR. Who Lives on the Wrong Side of the Environmental Tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model*. *Social Science Quarterly.* 2004;85(2):441-462.
35. *EPA's Risk-Screening Environmental Indicators (RSEI) Methodology: Version 2.3.9* 2020.
36. Mask D. Where the Streets Have No Name. *Atlantic.* 2014.
37. Ha S, Hu H, Mao L, Roussos-Ross D, Roth J, Xu X. Potential selection bias associated with using geocoded birth records for epidemiologic research. *Annals of epidemiology.* 2016;26(3):204-211.

38. Krieger N, Waterman P, Chen JT, Soobader M-J, Subramanian SV, Carson R. Zip code caveat: bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas--the Public Health Disparities Geocoding Project. *American journal of public health.* 2002;92(7):1100-1102.
39. Zandbergen PA. Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. *BMC Public Health.* 2007;7(1):37.
40. Cromley EK. Using GIS to Address Epidemiologic Research Questions. *Current Epidemiology Reports.* 2019;6(2):162-173.
41. Goldberg D, Swift, JN and Wilson JP. *Geocoding Best Practices: Reference Data, Input Data, and Feature Matching.* Los Angeles, CA, : University of Southern California GIS Research Laboratory Techincal Report No. 9;2008.
42. Krieger N, Waterman P, Lemieux K, Zierler S, Hogan JW. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. *American journal of public health.* 2001;91(7):1114-1116.
43. Buttling LG, McKnight MX, Kolivras KN, Ranganathan S, Gohlke JM. Maternal proximity to Central Appalachia surface mining and birth outcomes. *Environmental Epidemiology.* 2021;5(1):e128.
44. Jacquez GM. A research agenda: Does geocoding positional error matter in health GIS studies? *Spatial and Spatio-temporal Epidemiology.* 2012;3(1):7-16.
45. Andresen MA, Malleson N, Steenbeek W, Townsley M, Vandeviver C. Minimum geocoding match rates: an international study of the impact of data and areal unit sizes. *International Journal of Geographical Information Science.* 2020;34(7):1306-1322.
46. Schinasi LH, Auchincloss AH, Forrest CB, Diez Roux AV. Using electronic health record data for environmental and place based population health research: a systematic review. *Annals of epidemiology.* 2018;28(7):493-502.
47. Kotelchuck M. The Adequacy of Prenatal Care Utilization Index: its US distribution and association with low birthweight. *American journal of public health.* 1994;84(9):1486-1489.
48. Kind AJH, Buckingham WR. Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. *N Engl J Med.* 2018;378(26):2456-2458.
49. Health UoWSoMP. 2015 Area Deprivation Index v2.0. . In:2021.
50. Gilboa SM, Mendola P, Olshan AF, et al. Comparison of residential geocoding methods in population-based study of air quality and birth defects. *Environmental research.* 2006;101(2):256-262.
51. Oliver MN, Matthews KA, Siadaty M, Hauck FR, Pickle LW. Geographic bias related to geocoding in epidemiologic studies. *International journal of health geographics.* 2005;4(1):29.
52. Jones RR, DellaValle CT, Flory AR, et al. Accuracy of residential geocoding in the Agricultural Health Study. *International journal of health geographics.* 2014;13(1):37.
53. The New York Times. Kentucky Faces Tradition Of Offensive Road Names. In: The New York Times Company; 1996:13.
54. Paruk F, Moodley J. Maternal and neonatal outcome in early- and late-onset pre-eclampsia. *Seminars in Neonatology.* 2000;5(3):197-207.
55. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Practice & Research Clinical Obstetrics & Gynaecology.* 2011;25(4):391-403.
56. Raymond D, Peterson E. A Critical Review of Early-Onset and Late-Onset Preeclampsia. *Obstetrical & Gynecological Survey.* 2011;66(8).

57. Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *American journal of obstetrics and gynecology.* 2013;209(6):544.e541-544.e512.
58. Bicocca MJ, Mendez-Figueroa H, Chauhan SP, Sibai BM. Maternal Obesity and the Risk of Early-Onset and Late-Onset Hypertensive Disorders of Pregnancy. *Obstetrics & Gynecology.* 2020;136(1):118-127.
59. Staff AC. The two-stage placental model of preeclampsia: An update. *J Reprod Immunol.* 2019;134-135:1-10.
60. Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *American Journal of Obstetrics and Gynecology.* 2013;209(6):544.e541-544.e512.
61. Klonoff-Cohen HS, Savitz DA, Cefalo RC, McCann MF. An Epidemiologic Study of Contraception and Preeclampsia. *JAMA.* 1989;262(22):3143-3147.
62. Kho EM, McCowan LM, North RA, et al. Duration of sexual relationship and its effect on preeclampsia and small for gestational age perinatal outcome. *J Reprod Immunol.* 2009;82(1):66-73.
63. Cuschieri S. The STROBE guidelines. *Saudi journal of anaesthesia.* 2019;13(Suppl 1):S31-s34.
64. National Vital Statistics System (NVSS). Revisions of the U.S. Standard Certificates and Reports. 2019; https://www.cdc.gov/nchs/nvss/revisions-of-the-us-standard-certificates-and-reports.htm. Accessed June 13, 2021.
65. National Center for Health Statistics. Guide to completing the facility worksheets for the certificate of live birth and report of fetal death (2003 revision). . https://www.cdc.gov/nchs/data/dvs/GuidetoCompleteFacilityWks.pdf. Accessed Feb 3, 2021, 2021.
66. United States Department of Agriculture. Rural-Urban Continuum Codes. 2021; https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.
67. TIGER/Line® Shapefiles. In: United States Census Bureau, ed: United States Census Bureau 2016.
68. Appalachian Regional Commision (ARC). Counties in Appalachian 2020; https://www.arc.gov/appalachian_region/CountiesinAppalachia.asp. Accessed March 23, 2020, 2020.
69. Weir CBJ, A. BMI Classification Percentile And Cut Off Points. In: *StatPearls.* Treasure Island (FL): StatPearls Publishing; 2021.
70. England LJ, Levine RJ, Qian C, et al. Smoking before pregnancy and risk of gestational hypertension and preeclampsia. *American journal of obstetrics and gynecology.* 2002;186(5):1035-1040.
71. Sibai BM, Gordon T, Thom E, et al. Risk factors for preeclampsia in healthy nulliparous women: A prospective multicenter study. *American journal of obstetrics and gynecology.* 1995;172(2, Part 1):642-648.
72. Alexander GR, Kotelchuck M. Quantifying the adequacy of prenatal care: a comparison of indices. *Public Health Rep.* 1996;111(5):408-418; discussion 419.
73. Concept: Revised-Graduated Prenatal Care Utilization Index (R-GINDEX). In:2021.
74. Appalahcian Regional Commission A. Classifying Economic Distress in Appalachian Counties - Appalachian Regional Commission. In:2021.
75. Kulldorff M. SaTScan user guide for version 10.0. 2021; https://www.satscan.org/cgi-bin/satscan/register.pl/SaTScan_Users_Guide.pdf?todo=process_userguide_download. Accessed November 8, 2021.

76. Chen J, Yang S, Li H, Zhang B, Lv J. Research on geographical environment unit division based on the method of natural breaks (Jenks). *Int Arch Photogramm Remote Sens Spat Inf Sci.* 2013;3:47-50.
77. Joinpoint Regression Program, Version 4.8.0.1 - April 2020; Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute. In.
78. Anselin L, Rey SJ. *Modern spatial econometrics in practice : a guide to GeoDa, GeoDaSpace and PySAL.* 2014.
79. Schoenberg NE, Huang B, Seshadri S, Tucker TC. Trends in cigarette smoking and obesity in Appalachian Kentucky. *South Med J.* 2015;108(3):170-177.
80. Bonauto DK, Lu D, Fan ZJ. Obesity prevalence by occupation in Washington State, Behavioral Risk Factor Surveillance System. *Prev Chronic Dis.* 2014;11:130219-130219.
81. Wikstrom AK, Stephansson O, Cnattingius S. Tobacco use during pregnancy and preeclampsia risk: effects of cigarette smoking and snuff. *Hypertension (Dallas, Tex : 1979).* 2010;55(5):1254-1259.
82. Neels K, Murphy M, Ní Bhrolcháin M, Beaujouan É. Rising Educational Participation and the Trend to Later Childbearing. *Population and Development Review.* 2017;43(4):667-693.
83. Saftlas AF, Rubenstein L, Prater K, Harland KK, Field E, Triche EW. Cumulative exposure to paternal seminal fluid prior to conception and subsequent risk of preeclampsia. *Journal of Reproductive Immunology.* 2014;101-102:104-110.
84. Dekker G, Robillard PY, Roberts C. The etiology of preeclampsia: the role of the father. *J Reprod Immunol.* 2011;89(2):126-132.
85. Kahn LG, Trasande L. Environmental Toxicant Exposure and Hypertensive Disorders of Pregnancy: Recent Findings. *Current hypertension reports.* 2018;20(10):87.
86. Pollack AZ, Ranasinghe S, Sjaarda LA, Mumford SL. Cadmium and Reproductive Health in Women: A Systematic Review of the Epidemiologic Evidence. *Current Environmental Health Reports.* 2014;1(2):172-184.
87. Tellez-Plaza M, Jones MR, Dominguez-Lucas A, Guallar E, Navas-Acien A. Cadmium exposure and clinical cardiovascular disease: a systematic review. *Current atherosclerosis reports.* 2013;15(10):356.
88. Yazbeck C, Thiebaugeorges O, Moreau T, et al. Maternal blood lead levels and the risk of pregnancy-induced hypertension: the EDEN cohort study. *Environmental health perspectives.* 2009;117(10):1526-1530.
89. Miranda ML, Edwards SE, Chang HH, Auten RL. Proximity to roadways and pregnancy outcomes. *Journal of exposure science & environmental epidemiology.* 2013;23(1):32-38.
90. Salemi JL, Tanner JP, Sampat DP, et al. Evaluation of the sensitivity and accuracy of birth defects indicators on the 2003 revision of the US birth certificate: has data quality improved? *Paediatric and perinatal epidemiology.* 2017;31(1):67-75.
91. Faucett K, Kennedy HP. Accuracy in Reporting of Kentucky Certified Nurse-Midwives as Attendants in Birth Registration Data. *Journal of midwifery & women's health.* 2020;65(5):688-693.
92. Martin JA, Wilson EC, Osterman MJ, Saadi EW, Sutton SR, Hamilton BE. Assessing the quality of medical and health data from the 2003 birth certificate revision: results from two states. *National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System.* 2013;62(2):1-19.

93. Bell ML, Belanger K. Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. *Journal of exposure science & environmental epidemiology.* 2012;22(5):429-438.
94. Wier ML, Pearl M, Kharrazi M. Gestational age estimation on United States livebirth certificates: a historical overview. *Paediatric and Perinatal Epidemiology.* 2007;21(s2):4-12.
95. Zollinger T, Przybylski M, Gamache R. Reliability of Indiana birth certificate data compared to medical records. *Ann Epidemiol.* 2006;16 1:1-10.
96. DiGiuseppe DL, Aron DC, Ranbom L, Harper DL, Rosenthal GE. Reliability of birth certificate data: a multi-hospital comparison to medical records information. *Maternal and child health journal.* 2002;6(3):169-179.
97. Obesity in Pregnancy: ACOG Practice Bulletin, Number 230. *Obstet Gynecol.* 2021;137(6):e128-e144.
98. Chowdhury R, Ramond A, O'Keeffe LM, et al. Environmental toxic metal contaminants and risk of cardiovascular disease: systematic review and meta-analysis. *BMJ.* 2018;362:k3310.
99. Murrison LB, Brandt EB, Myers JB, Hershey GKK. Environmental exposures and mechanisms in allergy and asthma development. *The Journal of Clinical Investigation.* 2019;129(4):1504-1515.
100. Shen J, Liao Y, Hopper JL, Goldberg M, Santella RM, Terry MB. Dependence of cancer risk from environmental exposures on underlying genetic susceptibility: an illustration with polycyclic aromatic hydrocarbons and breast cancer. *British Journal of Cancer.* 2017;116(9):1229-1233.
101. Espinoza J. Low-Dose Aspirin for the Prevention of Preeclampsia. *JAMA.* 2021;326(12):1153-1155.
102. Institute of Medicine Committee on Assuring the Health of the Public in the 21st C. In: *The Future of the Public's Health in the 21st Century.* Washington (DC): National Academies Press (US)
Copyright 2003 by the National Academy of Sciences. All rights reserved.; 2002.
103. Trust for America's Health. *The state of obesity: Better policies for a healthier America.* 2021.
104. Kentucky's Environmental Public Health Tracking Network. Health Indicator Report of Tobacco Use - Adult Smoking Prevalence. 2021; https://kyibis.mc.uky.edu/ehl/dataportal/indicator/view/TobaccoSmokeAdult.Cnty.html . Accessed November 18, 2021.
105. Kentucky Geological Survey. Groundwater-Quality Analyte Descriptions. 2021; https://kgs.uky.edu/kgsweb/datasearching/water/analyteDescr.asp. Accessed October 9, 2021.
106. Laine JE, Ray P, Bodnar W, et al. Placental Cadmium Levels Are Associated with Increased Preeclampsia Risk. *PLoS One.* 2015;10(9):e0139341-e0139341.
107. University of Wisconsin School of Medicine Public Health. Area Deprivation Index 2019; https://www.neighborhoodatlas.medicine.wisc.edu/
108. Ways to Get RSEI Results. In: United States Environmental Protection Agency, ed. Agency USEP, trans: United States Environmental Protection Agency 2016.
109. United States Environmental Protection Agency. Basics of TRI Reporting. *US EPA* 2013; https://www.epa.gov/toxics-release-inventory-tri-program/basics-tri-reporting.

110. Tanwar V, Adelstein JM, Grimmer JA, et al. Preconception Exposure to Fine Particulate Matter Leads to Cardiac Dysfunction in Adult Male Offspring. *Journal of the American Heart Association.* 2018;7(24):e010797.
111. Harville EW, Mishra GD, Yeung E, et al. The Preconception Period analysis of Risks and Exposures Influencing health and Development (PrePARED) consortium. *Paediatric and perinatal epidemiology.* 2019;33(6):490-502.
112. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. *Structural equation modeling : a multidisciplinary journal.* 2007;14(4):671-694.
113. Lanza ST, Dziak, J. J., Huang, L., Wagner, A., & Collins, L. M, (2015). *PROC LCA & PROC LTA users' guide (Verision 1.3.2).* University Park: The Methodology Center: Penn State.
114. *Toxicological Profile for Cadmium* Agency for Toxic Substances & Disease Regsitry (ASTDR) ,;2012.
115. Poropat AE, Laidlaw MAS, Lanphear B, Ball A, Mielke HW. Blood lead and preeclampsia: A meta-analysis and review of implications. *Environmental research.* 2018;160:12-19.
116. Maduray K, Moodley J, Soobramoney C, Moodley R, Naicker T. Elemental analysis of serum and hair from pre-eclamptic South African women. *Journal of trace elements in medicine and biology : organ of the Society for Minerals and Trace Elements (GMS).* 2017;43:180-186.
117. Hill DT, Petroni M, Larsen DA, et al. Linking metal (Pb, Hg, Cd) industrial air pollution risk to blood metal levels and cardiovascular functioning and structure among children in Syracuse, NY. *Environmental research.* 2021;193:110557.
118. Ogneva-Himmelberger Y, Dahlberg T, Kelly K, Simas TAM. Using Geographic Information Science to Explore Associations between Air Pollution, Environmental Amenities, and Preterm Births. *AIMS Public Health.* 2015;2(3):469-486.
119. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prevention science : the official journal of the Society for Prevention Research.* 2013;14(2):157-168.
120. Stevens W, Shih T, Incerti D, et al. Short-term costs of preeclampsia to the United States health care system. *American journal of obstetrics and gynecology.* 2017;217(3):237-248.e216.
121. Khodzhaeva ZS, Kogan YA, Shmakov RG, et al. Clinical and pathogenetic features of early- and late-onset pre-eclampsia. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet.* 2016;29(18):2980-2986.
122. Han D, Bonner MR, Nie J, Freudenheim JL. Assessing bias associated with geocoding of historical residence in epidemiology research. *Geospatial health.* 2013;7(2):369-374.
123. Haidari ES, Lee HC, Illuzzi JL, Lin H, Xu X. Utility of Birth Certificate Data for Evaluating Hospital Variation in Admissions to NICUs. *Hospital Pediatrics.* 2020;10(2):190.

VITA

PLACE OF BIRTH – Corbin, KY

EDUCATION

Berea College, Berea KY

2013 BA: Psychology; Minor: Spanish

PROFESSIONAL EXPERIENCE

2017-Present	Graduate Research Assistant, Department of Epidemiology, University of Kentucky
2016-2014	Research Coordinator, University of Kentucky
2015-2014	Lab Manager, University of Kentucky

TEACHING EXPERIENCE

2018-2019	Adjunct Faculty, Department of Kinesiology and Health Science & Department of Mathematics, Georgetown College Georgetown, Kentucky
2015-2016,2018	Teaching Assistant, Department of Epidemiology, University of Kentucky, Lexington KY

SCHOLASTIC AND PROFESSIONAL HONORS

August 2017-Present	Central Appalachian Regional Education and Research Center (CAREC) Trainee - University of Kentucky
April 2021	Olive Ruth Russell Fellowship – Berea College ($2,000)
July 2020	Central Appalachian Regional Education and Research Center (CAREC) – Pilot Funding ($15,000)
April 2020	Appalachian Center Eller & Billings Student Research Award – University of Kentucky ($750)

PROFESSIONAL PUBLICATIONS

Walker, CJ, Levy, J., Christian, WJ, & Browning, SR (2021). Geocoding birth records from Kentucky Vital Statistics, 2008-2017. *Geospatial Health, Accepted*

Christian, WJ, Valvi, N., **Walker, CJ**., (2021*)* Comparing E-cigarette Use in Kentucky, its Appalachian region, and the United States using the Behavioral Risk Factor Surveillance System. *Public Health Reports.* doi:10.1177/00333549211029972

Walker, CJ & Christian, WJ & (2021*)* Estimating the population attributable fraction of asthma due to electronic cigarette use and other risk factors using Kentucky Behavioral Risk Factor Survey data, 2016-2017. *Substance Use and Misuse*, 56 (1), 1-6.

Christian WJ, **Walker, CJ**., Huang, B., Levy, JE., & Durbin, E (2020) Using residential histories in case-control analysis of lung cancer and mountaintop removal coal mining in Central Appalachia. *Spatio and Spatio-temporal Epidemiology*,35:100364

Christian WJ, **Walker, CJ**., Huang, B., & Hahn, EJ. (2019a) Effect of Local Smoke-Free Ordinances on Smoking Prevalence in Kentucky, 2002-2009.*Southern Medical Journal* 112,7

Christian, WJ. Vanderford, NL., McDowell, J., Huang, B., Durbin, EB., Absher, KJ., **Walker CJ.,** & Arnold, SM. (2019b) Spatiotemporal Analysis of Lung Cancer in Kentucky 1996-2014. *Cancer Control,* 26, 1:8