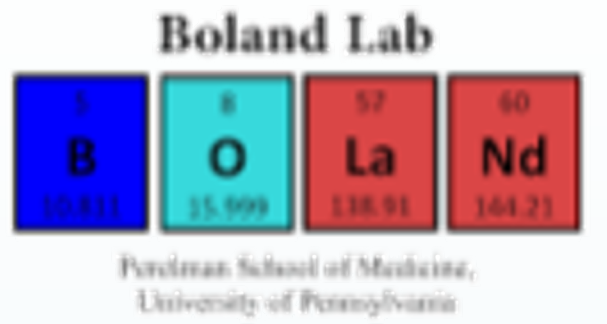


Incorporating Neighborhood-Level Covariates from Public Databases into a Geographic Analysis of Adverse Medical Outcomes in Philadelphia



Jessica Liu,¹ Mary Regina Boland², MA, MPhil, PhD, Ray Bai³, PhD, Cecilia Balocchi⁴, PhD, Silvia Canelón², PhD, Edward I. George⁴, PhD, Yong Chen², PhD.

¹Penn Undergraduate Research Mentoring Program, The Wharton School (Class of 2023), University of Pennsylvania

²Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania,

³Department of Statistics, University of South Carolina, ⁴Department of Statistics, The Wharton School, University of Pennsylvania

Background

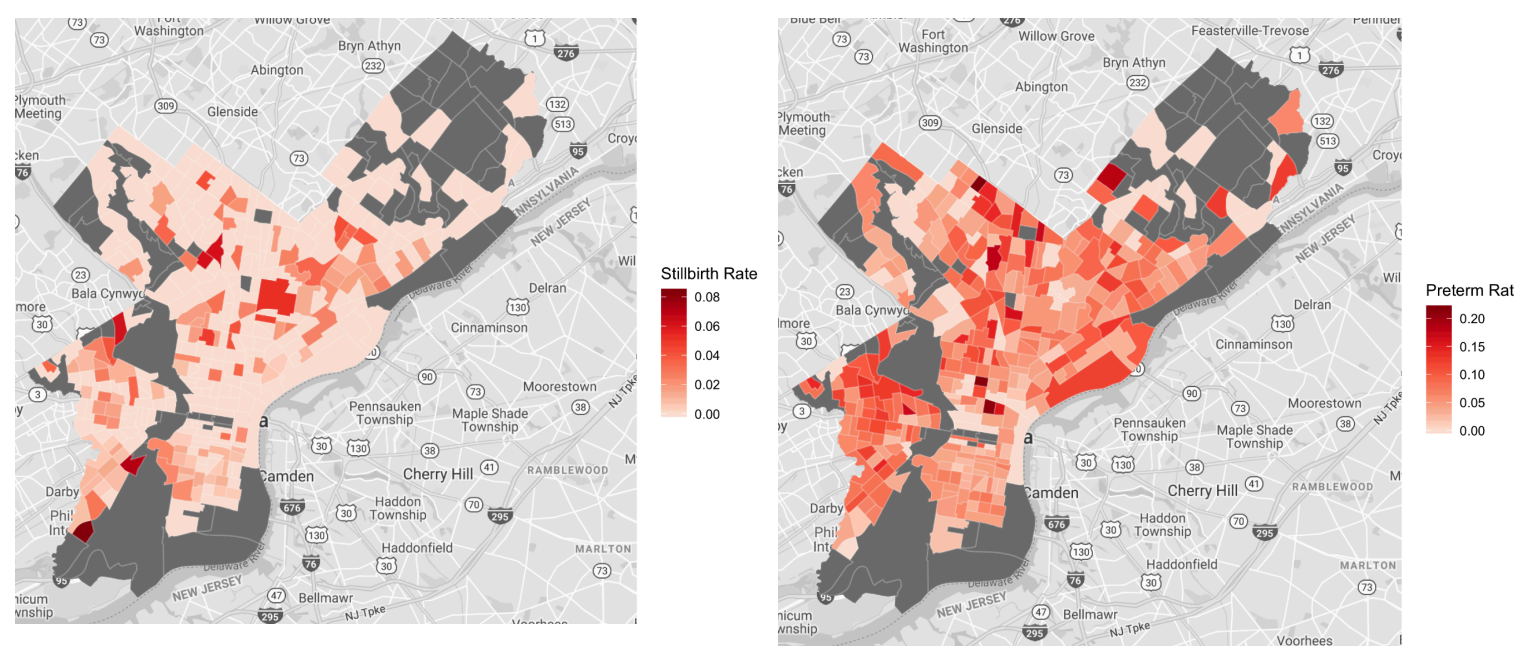
- Adverse pregnancy outcomes remain a major public health concern despite considerable advancements in maternal health over the recent decades.
- Stillbirth** is the delivery of a baby who has passed away, after the 20th week of pregnancy. Although it is rare (1/100 pregnancies), stillbirth can be devastating. A higher rate of stillbirth has been found among African-Americans, and it is believed that housing quality, air pollution, and various environmental risk factors influence these trends. However, there is no known correlation between ancestries and birth outcomes.
- Preterm birth** is when a baby is born before 37 weeks of pregnancy have been completed. This birth outcome is more common (1/10 of infants born in the U.S) and accounted for 17% of infant deaths in 2017. Some believed risk factors include housing issues (e.g. heavy metal exposures), stress, and obesity.
- The other birth outcomes investigated by this study were C-section rate and multiple birthrate. Over the course of the summer, the scope of the study was narrowed to focus on modeling stillbirth and preterm birth outcomes.

Goal: This study aims to determine whether geographic variation exists for patients who delivered babies at Penn Medicine and whether this geographic variation is correlated with certain birth outcomes. Using Electronic Health Records (EHR) from the Penn Medicine Medical Bank and publicly available neighborhood data at the census tract level, a model will be built to predict preterm birth and stillbirth in the city of Philadelphia.

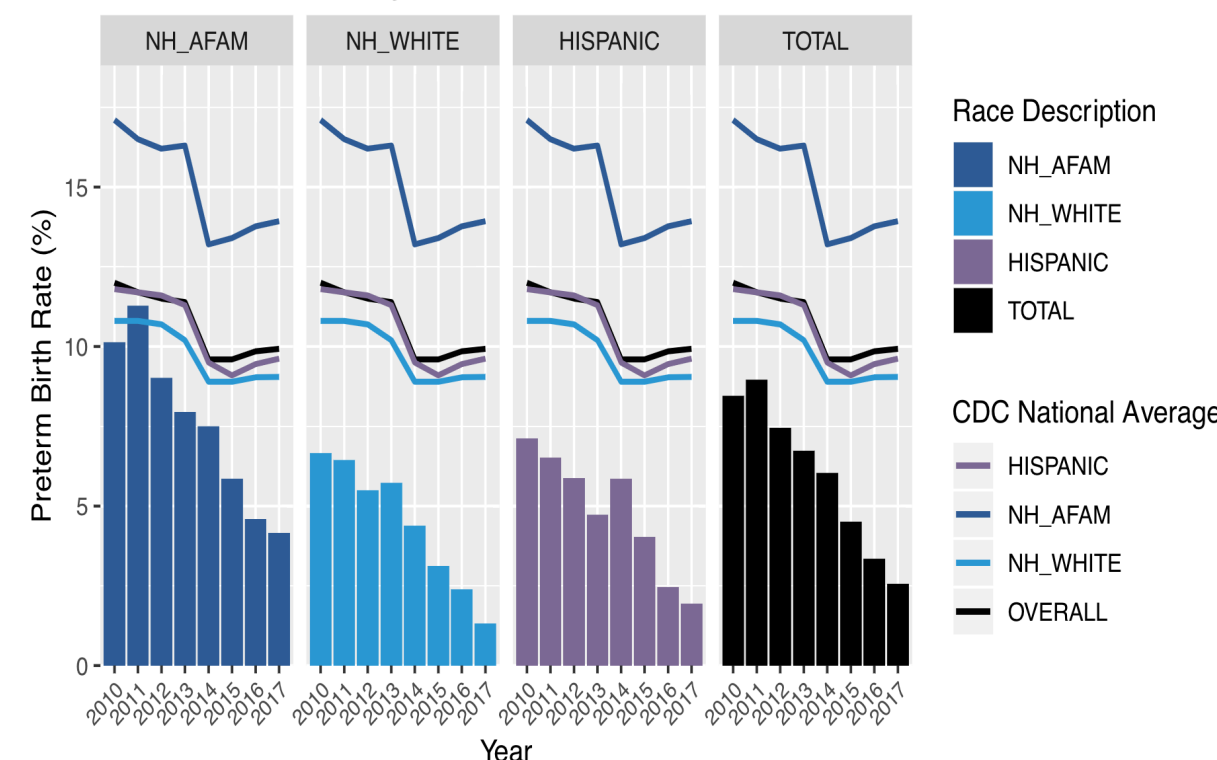
Previous Progress

Stillbirth Rate

Preterm Rate



Preterm Birth Rate by Year

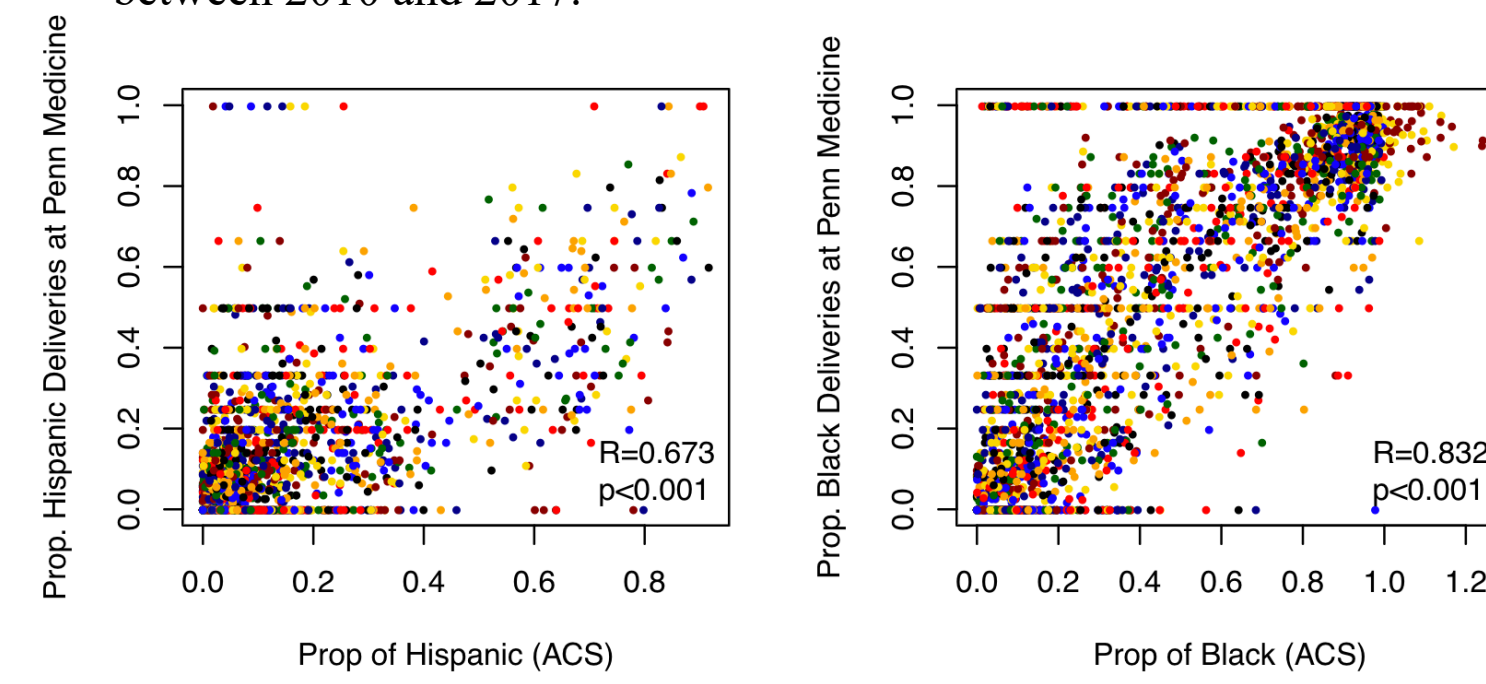


Objectives

- Use census data to add yearly level variation within current and newly added covariates to track and visualize changes within Philadelphia census tracts between 2010-2017.
- Determine the correlation of yearly neighborhood covariates to adverse medical outcomes.

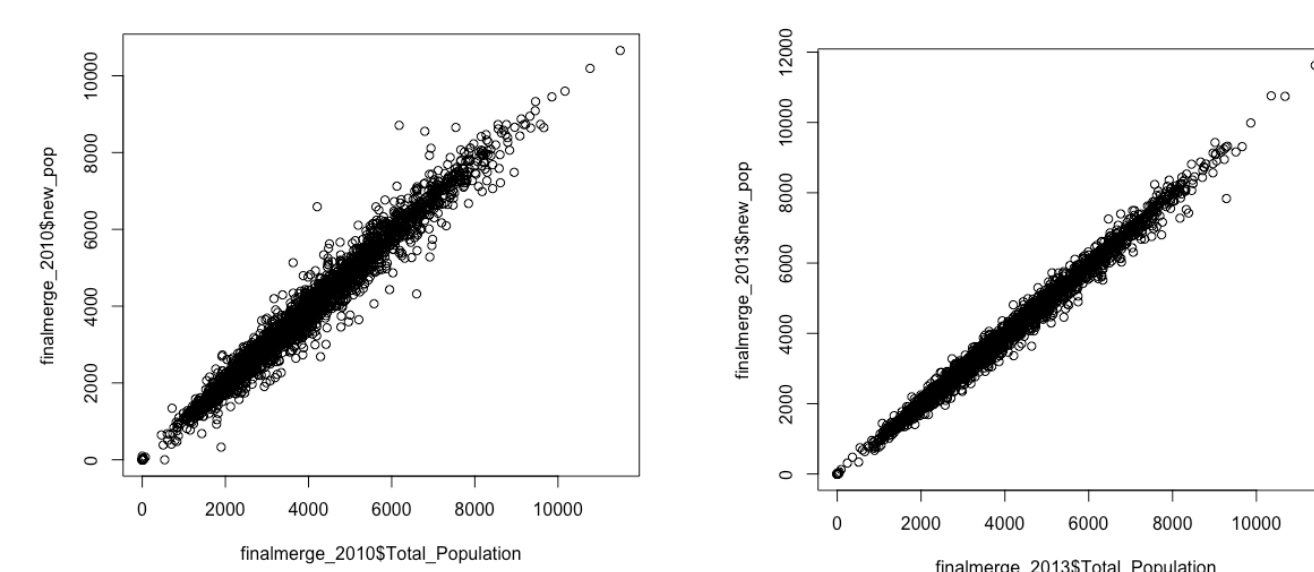
Data Methodology

- Raw data for yearly covariates for Pennsylvania on the census tract level was downloaded from the U.S. Census Bureau, data.census.gov/cedsci/.
- All tables utilized in the final covariate calculations came from the American Community Survey (ACS) 5-Year Estimates conducted between 2010 and 2017.



- Census tract population values were represented using the sum of the estimated totals of Asian Alone, Black or African-American Alone, White Alone, American Indian and Alaska Native Alone, Native and Other Pacific Islander Alone, and Some Other Race Alone from each respective census tract.

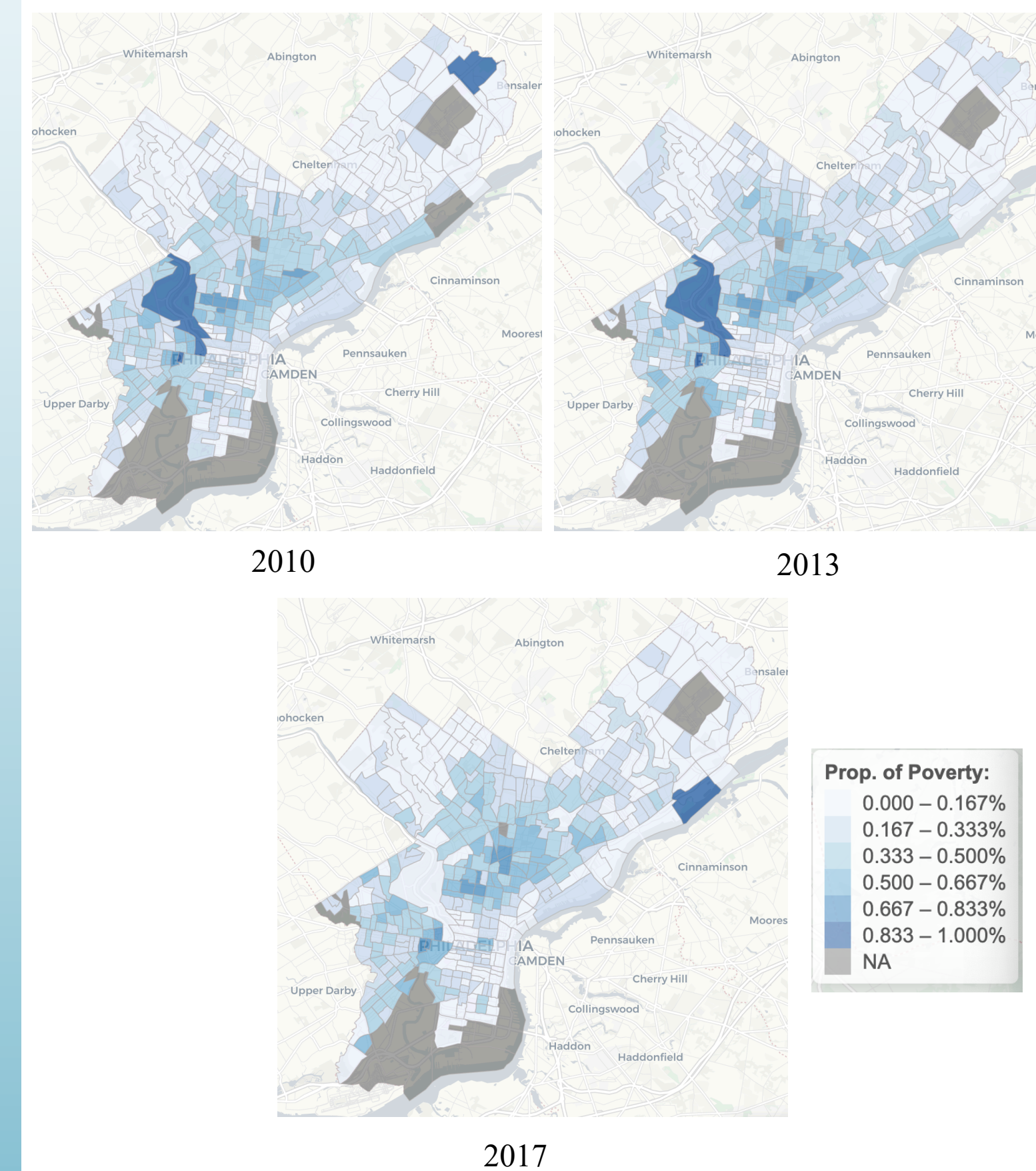
Representative Sum vs. ACS Population Estimates



- Covariates:** Proportion of Asian Alone, Proportion of Hispanic or Latinx Alone, Proportion of Black or African-American Alone, Proportion of White Alone, Median Income (Families), Proportion Owner-occupied Housing Units, Proportion Renter-occupied Housing Units, Percent Below Poverty Level (All Families), Proportion of Women 15-50 y/o, Proportion of Women 15-50 y/o with Poverty Status, Proportion of Women 15-50 y/o below 100% Poverty Level, Proportion of Women 15-50 y/o who Received Public Assistance in Past 12 Months, Proportion of Women 15-50 y/o with Births in the Past 12 Months, Proportion of Women 15-50 y/o with High School Diploma (or Equivalent), Proportion of Women 15-50 y/o with Bachelor's Degree

Preliminary Results

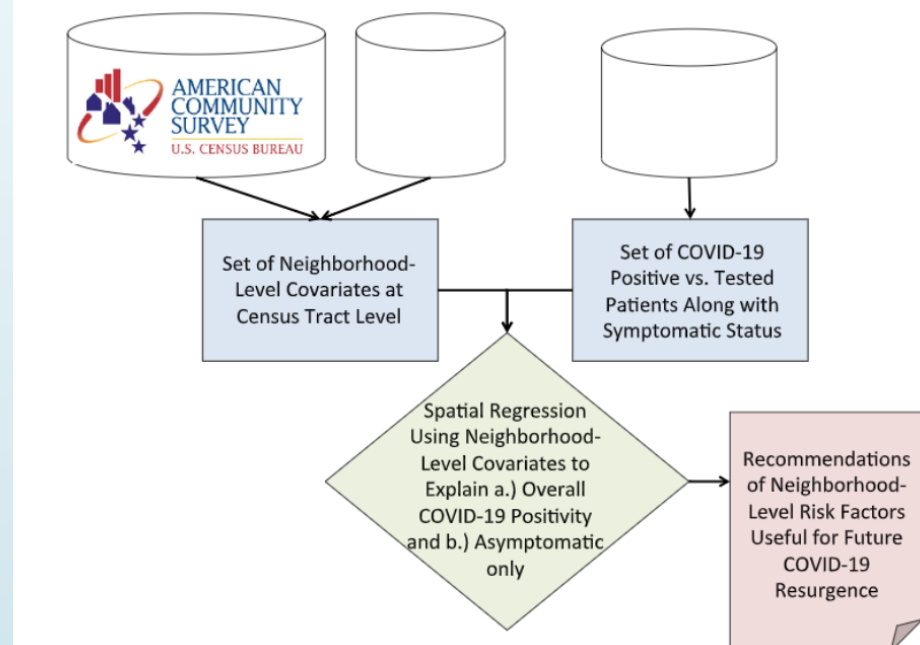
Example Maps: Proportion of Women 15-50 y/o with Poverty Status



Geospatial Analysis

- Preterm Birth:**
 - Avg. AUC = 0.59, but much better (0.68) for 2017 than other years.
 - Covariates significant ($p < 0.05$) in at least 50% of the training sets:
 - Prop. of Women 15-50 y/o
 - Prop. of Women 15-50 y/o with H.S. Diploma (or Equivalent)
 - Prop. of Women 15-50 with Bachelor's Degree.
- Stillbirth:**
 - Avg. AUC = 0.61, but much better (0.61) for 2017 than other years.
 - Covariates significant ($p < 0.05$) in at least 50% of the training sets:
 - Prop. of Hispanic
 - Prop. of Asian (neighborhood)
 - Prop. of Women 15-50 y/o w/Poverty Status
 - Prop. of Women 15-50 y/o who Received Public Assistance in Past 12 Months.

Application to COVID-19



- Philly neighborhood-level covariates were used in both univariate and multi-variate spatial regression models for COVID-19 positivity rates.

Table 4. Results of Multi-variate Spatial Regression Model on Neighborhood-Level Covariates for COVID-19 Positivity Rates: All Patients Regardless of Symptomology

Name	Neighborhood-Level Covariate	Odds Ratio	P-value
Education	Prop. of women aged 15-50 years in each census tract that graduated high school (including equivalency)	1.2114	<0.001
White	Prop. of each census tract that identifies as White Alone	0.8283	0.0011
Hispanic	Prop. of each census tract that identifies as Hispanic or Latinx	1.1302	0.0027
Public Assistance	Prop. of women aged 15-50 years in each census tract that received public assistance income in the past 12 months	0.8136	0.0406
Labor Force	Prop. of women aged 16-50 years in each census tract that are in the labor force	1.0942	0.1045
Black	Prop. of each census tract that identifies as Black or African-American Alone	0.9421	0.2708
Income	Median family income (dollars) (log-transformed variable)	0.9891	0.5592
Housing Violations	Housing Violations (log-transformed variable)	1.0027	0.7263
Violent Crime	Violent Crime Rate (log-transformed variable)	0.9989	0.9116
Poverty	Prop. of women aged 15-50 years in each census tract below 100 percent poverty level	1.0012	0.9816

- Positively associated with COVID-19: Education status (high-school), Hispanic/Latinx.
- Reduced risk of COVID-19: White, Received Public Assistance

Further Directions

- The utilization of yearly neighborhood-level covariates brings nuance into the building of predictive models for adverse pregnancy outcomes.
- Although not all training sets were significant, running the analysis for previous models after adding yearly neighborhood covariates revealed whether the consideration of yearly variation strengthened the predictive power of the model.
- If prediction results (AUC, sensitivity, specificity) do not improve after running the models for the full data set (all years), more focus will be put on investigating spatial trends rather than predictive accuracy.

References

S Canelón, HH Burris, LD Levine, MR Boland. 2020. Development and Evaluation of MADDIE: Method to Acquire Delivery Date Information from Electronic Health Records. *medRxiv*.

Acknowledgements

This work was made possible through funding from the Center for Undergraduate Research's Penn Undergraduate Research Mentoring (PURM) Program at the University of Pennsylvania.