NEIGHBORHOOD AND ENVIRONMENTAL PREDICTORS OF ASTHMA IN THE U.S.A.

By

COLLIN CHRISTOPHER ECKHAUSER

A Thesis Submitted to The W.A. Franke Honors College

In Partial Fulfillment of the Bachelors degree
With Honors in

Natural Resources

THE UNIVERSITY OF ARIZONA

M A Y 2 0 2 2

Approved by:

Dr. Chris Lim
Department of Public Health

Neighborhood and Environmental Predictors of Asthma in the U.S.A.

By

COLLIN CHRISTOPHER ECKHAUSER

Abstract

Over the next several decades the United States will be severely affected by climate change. Consequently, the public health of the country and its human inhabitants and communities will be compromised. Select populations such as children, the elderly, and those living in squalid conditions or disadvantaged communities, will be disproportionately vulnerable. Asthma is a disease with a complex etiology, air pollution exacerbates symptoms of asthma and can cause the onset of asthma in adults. Asthma, like climate change, disproportionately affects certain groups of people. We created a dataset at the census tract level, by collecting data from multitudes of sources on environmental quality, behavioral factors, socioeconomic status, educational attainment, neighborhood characteristics, modifiable risk factors, and comorbidities of asthma. We created a correlation matrix to analyze which variables are positively and negatively correlated with asthma prevalence. Furthermore, we created multivariate linear regression models with our variables to determine the associations between asthma prevalence and neighborhood and environmental predictors. Neighborhoods with a higher percentage of Black people are positively correlated with higher asthma prevalence with a value of 0.50. Moreover, the highest positive correlation existed between es†nytimated percent of adults ever diagnosed with heart disease and estimated percent of adults ever diagnosed with COPD, emphysema, and chronic bronchitis in 2018, with a value of 0.90. In our regression models we

found that for a one unit increase of percent Black people in a neighborhood, there is an increase of 1.98 units of asthma prevalence. We also found in our models that for a one unit increase in estimated percent of adults reporting 14 or more days of poor physical health in the past 30 days in 2018, there is an increase of 0.210 units of asthma prevalence. By cultivating a greater understanding of how these factors interact simultaneously the strategic optimization of neighborhood level interventions can be implemented for asthma prevention.

Introduction

More Americans than ever before have asthma, it is one of this country's most common and costly diseases. Asthma is a chronic respiratory disease with a complex etiology that can be fatal. Over 25 million Americans have asthma, with approximately 20 million adults over the age of 18, and 5.1 million children under the age of 18 having asthma [1]. Asthma is more common in adult women than adult men, and it is more common in boys than girls in the United States. Furthermore, asthma disproportionately affects Black Americans, as they are nearly three times more likely to die from asthma than white Americans [1].

Over the next several decades the United States will be severely affected by climate change. Consequently, the public health of the country and its human inhabitants and communities will be compromised. Select populations such as children, the elderly, and those living in squalid conditions or disadvantaged communities, will be disproportionately vulnerable. Disadvantaged communities are more vulnerable to climate change because they are the least able to prepare for and recover from heat waves, poor air quality, and other impacts [5]. Children are disproportionately vulnerable to air pollution because it may disrupt the proper development of the lungs. Moreover, the elderly are disproportionately vulnerable to air pollution because the

ability to eliminate chemicals from the body decreases with age [9]. Climate change deteriorates the sustainability of life by inducing changes in the atmosphere mainly resulting from the buildup of greenhouse gasses from burning fossil fuels to provide energy for human activities, these noxious environmental changes can harm people's health and wellbeing, particularly respiratory functioning. Climate change can influence respiratory health by directly promoting or aggravating respiratory diseases or by increasing exposure to risk factors for respiratory diseases [3]. Air pollution and climate change are inextricably linked. The most severe respiratory health effects of air pollution are attributable to particulate matter and ozone, according to epidemiological studies [3]. Previous studies have also suggested that exposure to air pollution can lead to the development as well as exacerbation of asthma [4].

Asthma is influenced by biological, social, and environmental factors throughout the course of an individual's lifetime. However, asthma disproportionately impacts communities of color and low socioeconomic status [14]. There is a one-and-half-fold increase in the prevalence of asthma among individuals of low socioeconomic status (SES) compared with those of high SES. Moreover, extensive research conducted over the past decades support that low SES and environmental exposures connected to poverty are substantially associated with the development of asthma, and poor asthma outcomes [14]. Behavioral factors such as sedentary lifestyle, smoking, unhealthy diets, and obesity in individuals with asthma have been found to have increased negative effects from pollutant exposure [11]. Communities with low SES are characterized by low-income, low educational attainment, lower quality home environments, increased stress, polluted air and water, and high crime rates. These factors perpetuate a cycle of poor health management and intervention in low SES neighborhoods which leaves these

communities to be disproportionately affected by asthma in terms of prevalence rates and disease morbidity [11].

The complex nature of asthma leads it to be influenced by a variety of risk factors. These factors are non-modifiable (genetics and sex), and modifiable (environment and behavior) risks. Current evidence suggests causal roles for modifiable exposures on asthma such as obesity, unhealthy diet, and indoor or outdoor pollutants [15]. Furthermore, exposure to land use contaminants from living near an industrial park have been associated with increased risk of asthma in adults [6]. Both environmental and behavioral factors have been associated with increased asthmatic symptoms [8].

The goal of this paper is to substantiate public health research on many different parts of the neighborhood environment, with a focus on air pollution as it correlates to increased asthma incidence on a neighborhood scale. This is accomplished in two parts. We first created a dataset of census tract-level factors potentially associated with increased asthma symptoms, with an emphasis on socioeconomic disparities, neighborhood demographics, and neighborhood proximity to polluting sources. Using this dataset, we then applied advanced statistical and spatial modeling in RStudio to illustrate which factors are most highly associated with rates of asthma incidence and where these increased rates of asthma occur spatially across the nation. In this part, we developed a multivariate linear regression model to examine a variety of factors associated with asthma while making considerations for air pollution on a neighborhood scale. Commonly in the United States, minority populations are concentrated in neighborhoods with a low overall socioeconomic status. We define socioeconomic status in this paper as socioeconomic status and social vulnerability level. Moreover, we use factors such as estimated percent of all people living in poverty, estimated per capita income, estimated percent of all

households that rent a home or own a home, estimated percent of people with a high school diploma, and predominant racial or ethnic group. We are interested in determining if low socioeconomic status as we define it and in terms of our considerations for correlation factors listed previously on a neighborhood scale shows that these neighborhoods are disproportionately vulnerable to the effects of climate change, namely air pollution.

Analyzing the impact on asthma prevalence on a neighborhood scale by a variety of elements related to environmental quality, behavioral factors, socioeconomic status, educational attainment, neighborhood characteristics, modifiable risk factors, and comorbidities can facilitate a greater understanding of how these factors interact simultaneously with the complex exposures to air pollution and how this affects asthma prevalence. This research paper aims to ultimately identify areas that are disproportionately impacted by climate change with an emphasis on environmental justice in public health.

Methods

The data used for this research was collected at the census tract level. The source for the data collection at the census tract level was PolicyMap. PolicyMap is an interactive tool that allows you to set geographical boundaries and define variables of interest to gather data on the defined areas in terms of the variables of interest. The variables of interest were selected on the basis of prior evidence from studies analyzing environmental predictors of asthma, and variables were selected due to their nature as a risk factor of asthma. We used this tool due to the extensive dataset that is available and its accessibility. Furthermore, we were able to analyze demographics, incomes and spending, housing, quality of life, economy, education, and health through PolicyMap [12]. This facilitated the means for us to create a comprehensive dataset

encompassing the entirety of the United States at the census tract level that included variables of interest for the purposes of the study. These variables of interest include estimated percent of adults reporting to have asthma in 2018 (eparha18), predominant racial or ethnic group between 2013-2017 (preg13_17), estimated percent of all people living in poverty as of 2015-2019 (epaplp15_19), estimated per capita income between 2015-2019 (epci15_19), estimated percent of people with a high school diploma (eppwhsd), estimated percent of all households that rent a home or own a home between 2015-2019 (epahr15_19 & epaho15_19), socioeconomic status social vulnerability level as of 2018 (sessv18), social vulnerability level as of 2018 (svl18), estimated median age of all people between 2015-2019 (emaap15_19), estimated percent of adults ever diagnosed with hypertension (epaedhbp), estimated percent of adults reporting to be obese (BMI >30) (eparo), estimated percent of adults reporting to be physically inactive in the past 30 days (eparpip_30d), estimated percent of adults ever diagnosed with hyperlipidemia (epaedhc), estimated percent of adults ever diagnosed with heart disease (epaedhd), estimated ratio of men to women (number of males per hundred females) (ermtw_nm_per100f), estimated percent of adults ever diagnosed with COPD, emphysema, or chronic bronchitis in 2018 (epaed_copd_emphy_cb_18), estimated percent of adults reporting to have ever smoked cigarettes (eparhesc), life expectancy at birth as of 2010-2015 (leab10_15), cooling degree days 2008 (cddr08), heating degree days 2008 (hddr08), estimated percent of adults reporting 14 or more days of poor physical health in the past 30 days in 2018 (epar14_ormdpph_p30d_18), and estimated percent of adults ever diagnosed with depression in 2018 (epaedd18), . Each one of these factors was included in a dataset of 72371 census tract levels.

PolicyMap served as the centralized source from which we gathered our data from. The original source of the data for a variety of variables utilized in the study is the Behavioral Risk

Factor Surveillance System (BRFSS). The BRFSS is the nation's premier health-related telephone survey that collects data across all 50 states about U.S. residents regarding their health-related behaviors, chronic conditions, and use of preventive services [2]. The variables used in the study from the BRFSS include estimated percent of adults reporting to have asthma in 2018 (eparha18), estimated percent of adults ever diagnosed with hypertension (epaedhbp), estimated percent of adults reporting to be obese (BMI >30) (eparo), estimated percent of adults reporting to be physically inactive in the past 30 days (eparpip_30d), estimated percent of adults ever diagnosed with hyperlipidemia (epaedhc), estimated percent of adults ever diagnosed with heart disease (epaedhd), estimated percent of adults ever diagnosed with COPD, emphysema, or chronic bronchitis in 2018 (epaed_copd_emphy_cb_18), estimated percent of adults reporting to have ever smoked cigarettes (eparhesc), life expectancy at birth as of 2010-2015 (leab10_15), cooling degree days 2008 (cddr08), heating degree days 2008 (hddr08), estimated percent of adults reporting 14 or more days of poor physical health in the past 30 days in 2018 (epar14_ormdpph_p30d_18), and estimated percent of adults ever diagnosed with depression in 2018 (epaedd18).

The original source of the remainder of variables utilized in this study are derived from the United States Census and the American Community Survey (ACS). The variables derived from the census and ACS include predominant racial or ethnic group between 2013-2017 (preg13_17), estimated percent of all people living in poverty as of 2015-2019 (epaplp15_19), estimated per capita income between 2015-2019 (epci15_19), estimated percent of people with a high school diploma (eppwhsd), estimated percent of all households that rent a home or own a home between 2015-2019 (epahr15_19 & epaho15_19), socioeconomic status social vulnerability level as of 2018 (sessv18), social vulnerability level as of 2018 (svl18), estimated

median age of all people between 2015-2019 (emaap15_19), and estimated ratio of men to women (number of males per hundred females) (ermtw_nm_per100f).

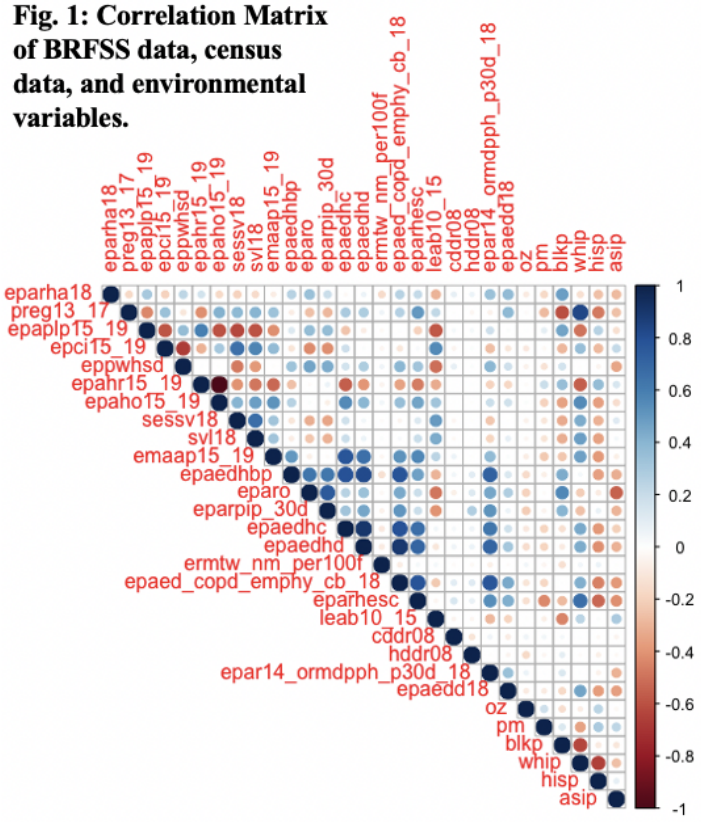
The air quality data we used in this study is data on particulate matter (PM), and ozone. This data was collected from the EPA and merged with our census tract dataset. Annual O3 concentrations were assigned to each participant's residence census tract using the EPA Fused Air Quality Surface Using Downscaling data, which is available for the years after 2002 [10]. A Bayesian space-time downscaler model is used to "fuse" daily 8-hour maximum O3 concentrations at the census tract centroid. The downscaler model develops a relationship between observed and modeled concentrations, and uses that relationship to spatially predict measurements at new locations in the spatial domain based on input data. Estimates are based on National Air Monitoring Stations/State and Local Air Monitoring Stations and Model-3/Community Multiscale Air Quality model data in 12 × 12-km grids. We aggregated daily data to annual averages [10].

We used our dataset to create a correlation matrix to illustrate the correlation of asthma prevalence with the variety of other factors in our dataset. Moreover, we developed a linear regression model to most effectively integrate a variety of factors associated with asthma while making considerations for air pollution on a neighborhood scale. Furthermore, we used our dataset to develop maps to show the distribution of variables across the United States.

Results

The results from the creation of our correlation matrix of all variables used in our study illustrate how asthma and other variables are correlated with each other, in some cases positively correlated and in other cases negatively correlated. There is a positive correlation of 0.5 between estimated percent of adults reporting to have asthma in 2018 and neighborhood makeup in terms of percent black people, which can be seen in Fig. 1. Moreover, there is a negative correlation of -0.6 between estimated percent of all people living in poverty as of 2015-2019 with estimated per capita income between 2015-2019. A positive correlation of 0.6 was found between estimated percent of all people living in poverty as of 2015-2019 and estimated percent of all households that rent a home between 2015-2019. Comparatively, we found that there was a negative correlation of -0.7 between estimated percent of all people living in poverty as of 2015-2019 with estimated percent of all households that own a home between 2015-2019. Estimated percent of all people living in poverty as of 2015-2019 was negatively correlated with life expectancy at birth as of 2010-2015, with a value of -0.6.

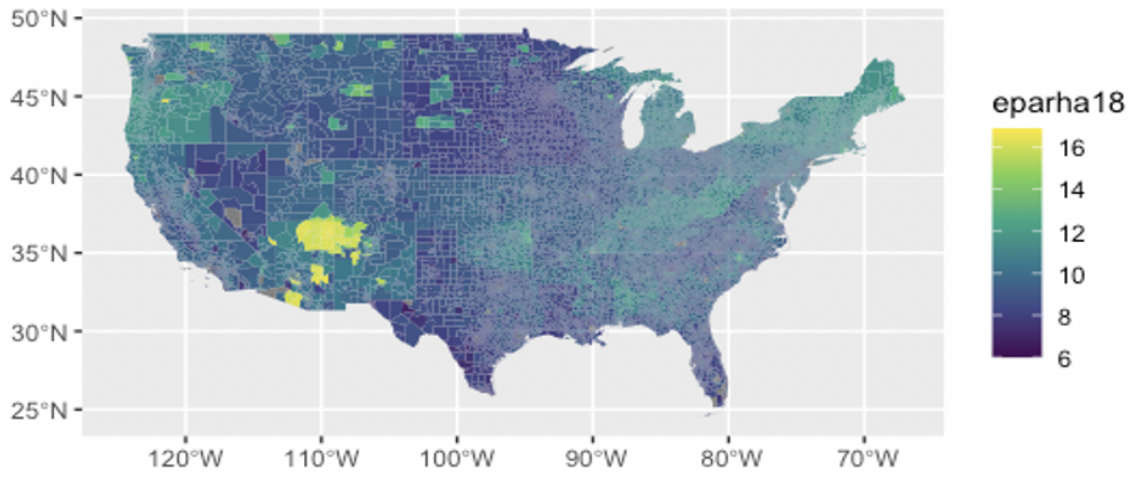
Fig. 1: Correlation Matrix of BRFSS data, census data, and environmental variables.



We found that there was a positive correlation of 0.6 between estimated percent of adults ever diagnosed with high blood pressure or hypertension and estimated percent of adults reporting to be obese (a body mass index of 30 or greater). Similarly, we found a positive correlation of 0.6 between estimated percent of adults ever diagnosed with high blood pressure and estimated percent of adults reporting to be physically inactive in the past 30 days. There is a positive correlation of 0.8 between estimated percent of adults ever diagnosed with high blood pressure and estimated percent of adults ever diagnosed with high cholesterol or hyperlipidemia. In the same respect, a positive correlation of 0.8 exists between estimated percent of adults ever diagnosed with high blood pressure and estimated percent of adults ever diagnosed with heart disease. A positive correlation of 0.8 exists between estimated percent of adults ever diagnosed with high blood pressure and estimated percent of adults ever diagnosed with chronic obstructive pulmonary disease (COPD), emphysema, and chronic bronchitis in 2018.

Furthermore, there is a positive correlation of 0.7 between estimated percent of adults ever diagnosed with high blood pressure and estimated percent of adults reporting 14 or more days of poor physical health in the past 30 days in 2018. We also found a positive correlation of 0.9 between estimated percent of adults ever diagnosed with heart disease and estimated percent of adults ever diagnosed with COPD, emphysema, and chronic bronchitis in 2018.

Fig. 2: Map of asthma rate in U.S.A.



The mapping of estimated percent of adults reporting to have asthma in 2018 enabled us to visualize the prevalence of asthma rates across the country. As illustrated in Fig. 2, the highest rates of asthma exist in the Four Corners region which includes Arizona, Utah, Colorado, and New Mexico. The Four Corners region is surrounded by Native American reservations, and is home to the Navajo Nation, which is the largest Native American reservation in the United States. Furthermore, the highest rates of asthma in the country exist in Arizona. The high rates of asthma that exist in Arizona, which can be seen in Fig. 2, illustrate a stark contrast of asthma rates across the nation. No other state in the U.S.A has asthma rates as high as Arizona, or across

Fig. 3: Map of socioeconomic status social vulnerability level as of 2018 (sessv18) in the U.S.A.

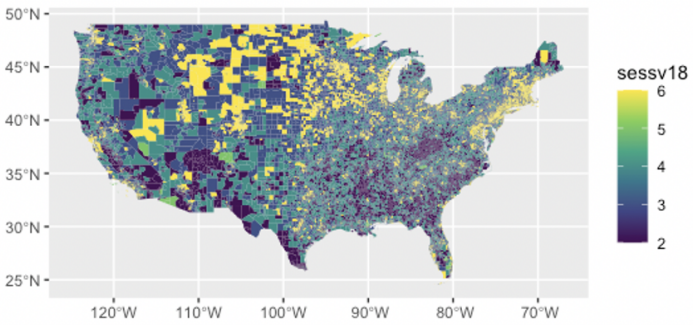
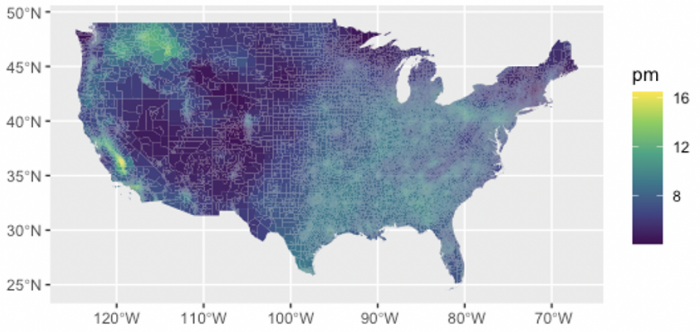


Fig. 4: Map of particulate matter (pm) levels across the U.S.A.



such a great area. This could be due to people moving to the desert hoping to avoid developing asthma, or alleviate respiratory illnesses [13].

We also mapped the socioeconomic status social vulnerability level as of 2018 at the census tract level across the country in Fig. 3. The highest rates of socioeconomic status social vulnerability level as of 2018 is seen in the northern states of the U.S.A. States such as North Dakota, Minnesota, South Dakota, Wyoming, and Nebraska experience the highest rates of socioeconomic status social vulnerability level. Moreover, we mapped the concentration of particulate matter levels by census tract across the country. The highest concentrations of particulate matter exist in California, Washington, and a moderate concentration across the midwest and south. This could be due to the large metropolitan complexes in California, and the high rate of wildfires in the region. The midwest and south could have a moderate concentration of particulate matter due to the large amounts of industry in the region.

Table 1 represents a summary of the minimum, maximum, median, mean, standard deviation, and interquartile range of the variables used in the study. We also included a column detailing the number of not applicable (NA) for the variety of variables. The NA's represent missing data at the census tract level.

Table 2 contains the findings from our variable selection for our linear regression modeling. We identified the variables that held significance from our all variable linear

Table 1: Summary statistics of variables used in the study.

Variable	Min	1st Qu	Median	Mean	3rd Qu	Max	NA's	Std
eparha18	6.000	8.900	9.600	9.703	10.400	16.900	743	1.233024
preg13_17	2.000	14.000	23.000	20.320	25.000	25.000	NA	6.466974
epaplp15_19	0.000	6.260	11.490	14.660	19.790	100.000	806	11.55785
epci15_19	14.000	22675.000	29714.000	33647.000	39774.000	279334.000	403	17548.62
eppwhsd	0.000	19.940	28.420	27.810	35.860	100.000	677	11.15476
epahr15_19	0.000	18.330	31.450	36.490	51.100	100.000	619	22.88347
epaho15_19	0.000	48.900	68.550	63.510	81.670	100.000	619	22.88347
sessv18	2.000	3.000	4.000	3.757	5.000	6.000	NA	1.476794
svl18	2.000	3.000	4.000	3.762	5.000	6.000	NA	1.477996
emaap15_19	6.000	34.000	39.000	39.680	44.000	83.000	385	7.838365
epaedhbp	5.700	28.100	32.000	32.250	36.100	60.900	691	6.60899
eparo	12.100	29.100	31.800	31.880	34.900	47.100	686	4.546873
eparpip_30d	13.600	26.300	30.000	30.160	34.200	49.700	693	5.606636
epaedhc	6.600	27.500	30.800	30.440	33.600	52.700	693	4.935703
epaedhd	0.200	3.100	4.000	4.082	4.900	13.900	947	1.406497
ermtw_nm_per100f	0.000	88.000	97.000	104.900	105.000	90250.000	474	451.9904
epaed_copd_emphy_cb_18	1.300	5.500	7.000	7.075	8.500	16.800	769	2.114862
eparhesc	11.300	37.000	41.300	40.780	44.900	59.300	676	5.500528
leab10_15	56.300	75.800	78.500	78.300	81.000	97.500	5551	3.985064
cddr08	1.000	4.000	5.000	5.180	6.000	11.000	NA	2.28845
hddr08	1.000	3.000	6.000	5.477	8.000	11.000	NA	2.51204
epar14_ormdpph_p30d_18	4.200	11.700	13.000	13.000	14.300	25.100	461	2.169923
epaedd18	9.300	17.500	19.300	19.690	21.500	32.900	437	3.050498
oz	27.85	36.46	37.92	38.55	39.43	59.05	NA	4.031311
pm	4.005	7.661	8.529	8.490	9.269	16.460	NA	1.493331
blkp	0.000	0.009928	0.037376	0.134043	0.14183	1.000	NA	0.2192576
whip	0.000	0.4345	0.7419	0.6413	0.8943	1.000	NA	0.3003646
hisp	0.000	0.02408	0.06204	0.15343	0.18251	1.000	NA	0.2090283
asip	0.000	0.004546	0.014161	0.041972	0.041654	1.000	NA	0.07956392

regression model. We selected these variables to be used in our adjusted variable linear regression model, the results of which can be found in Table 4.

Table 3 contains the coefficient values, standard errors, and P-values from our all variable linear regression model. The all variable linear regression model has an adjusted R-squared value of 0.6585, and a multiple R-squared value of 0.6589. We used the all variable linear regression

model to identify variables of significance, to be used in our adjusted variable linear regression model.

Table 4 represents the results of our adjusted variable linear regression model. We used our all variable linear regression model to identify variables that are significant to be used in our adjusted variable linear regression model. The adjusted variable linear regression model has an adjusted R-squared value of 0.6284, and a multiple R-squared value of 0.6286. Table 5 represents the results from our revised adjusted linear regression model. This model is the same as the adjusted model, except for the factors included. We substituted preg13_17 for blkp, whip, hisp, and asip to clarify the predominant racial and ethnic groups in a neighborhood. Moreover, we eliminated the factors cddr08 and hddr08 to simplify

Table 2: Variables of significance from all variable linear regression model.

	Estimate	Standard Error	P Value
epaplp15_19	5.81E-03	4.85E-04	< 2.00E-16***
epci15_19	7.60E-06	3.07E-07	< 2.00E-16***
eppwhsd	7.99E-03	4.30E-04	< 2.00E-16***
emaap15_19	7.33E-03	7.45E-04	< 2.00E-16***
epaedhbp	5.46E-02	2.46E-03	< 2.00E-16***
eparo	-4.80E-02	1.66E-03	< 2.00E-16***
eparpip_30d	6.39E-03	1.12E-03	1.29E-08***
epaedhc	-7.09E-02	2.70E-03	< 2.00E-16***
epaedhd	-4.52E-01	8.42E-03	< 2.00E-16***
ermtw_nm_per100f	-6.04E-03	1.53E-04	< 2.00E-16***
epaed_copd_emphy_cb_18	5.32E-02	5.06E-03	< 2.00E-16***
eparhesc	5.25E-02	1.37E-03	< 2.00E-16***
leab10_15	1.82E-02	1.03E-03	< 2.00E-16***
epar14_ormdpph_p30d_18	1.86E-01	3.74E-03	< 2.00E-16***
epaedd18	1.66E-01	1.61E-03	< 2.00E-16***
pm	-1.52E-01	2.55E-03	< 2.00E-16***
oz	-7.63E-03	8.44E-04	< 2.00E-16***
factor(preg13_17)3	-1.61E+00	7.13E-01	0.02362*
factor(preg13_17)4	-4.53E-01	4.68E-02	< 2.00E-16***
factor(preg13_17)5	-7.43E-01	7.55E-02	< 2.00E-16***
factor(preg13_17)6	9.51E-01	3.80E-02	< 2.00E-16***
factor(preg13_17)7	2.72E+00	4.48E-02	< 2.00E-16***
factor(preg13_17)8	1.43E+00	3.72E-02	< 2.00E-16***
factor(preg13_17)9	2.27E+00	4.05E-02	< 2.00E-16***
factor(preg13_17)10	2.42E-01	3.43E-02	1.56E-12***
factor(preg13_17)11	-4.54E-01	4.11E-02	< 2.00E-16***
factor(preg13_17)12	1.60E-01	3.35E-02	1.80E-06***
factor(preg13_17)15	1.24E+00	1.44E-01	< 2.00E-16***
factor(preg13_17)16	2.94E+00	1.07E-01	< 2.00E-16***
factor(preg13_17)17	1.30E+00	1.26E-01	< 2.00E-16***
factor(preg13_17)18	2.17E+00	1.09E-01	< 2.00E-16***
factor(preg13_17)19	2.02E+00	3.20E-01	2.51E-10***
factor(preg13_17)20	5.09E-01	1.52E-01	0.000778***
factor(preg13_17)22	2.13E-01	3.27E-02	8.10E-11***
factor(preg13_17)23	8.43E-02	3.40E-02	0.013207*
factor(preg13_17)24	1.30E-01	3.19E-02	4.53E-05***
factor(preg13_17)25	7.31E-02	3.26E-02	0.025077*
factor(svi18)3	-6.84E-02	1.38E-02	7.13E-07***
factor(svi18)4	-5.80E-02	1.06E-02	4.60E-08***
factor(svi18)6	-3.60E-02	1.71E-02	0.034885*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

the model. The revised adjusted linear regression model has a multiple R-squared value of 0.6675, and an adjusted R-squared value of 0.6673. The interpretation of the coefficients in our models is described in the following examples. For our revised adjusted linear regression model in Table 5, a one unit increase in eparo will result in a decrease of 0.07 units of estimated percent of adults reporting to have asthma in 2018. This decrease is because the coefficient of eparo is negative. A positive coefficient example in Table 5 is eparhesc, for a one unit increase in eparhesc there will be an increase of 0.0453 units of estimated percent of adults reporting to have asthma in 2018. The R-squared value for all three models in our study falls in between 0.50 and 0.70. This means that our model has a moderate fit. In other words, the movements or variation of our dependent variables are moderately explained by the movements of our independent variable.

Table 3: Coefficients, Standard Error, and P-Value results from all variable linear regression model.

	Estimate	Standard Error	P Value	
(Intercept)	-5.45E+02	9.05E+02		0.546692
epaplp15_19	5.81E-03	4.85E-04	< 2.00E-16	***
epci15_19	7.60E-06	3.07E-07	< 2.00E-16	***
eppwhsd	7.99E-03	4.30E-04	< 2.00E-16	***
epahr15_19	5.50E+00	9.05E+00		0.543112
epaho15_19	5.50E+00	9.05E+00		0.543557
emaap15_19	7.33E-03	7.45E-04	< 2.00E-16	***
epaedhbp	5.46E-02	2.46E-03	< 2.00E-16	***
eparo	-4.80E-02	1.66E-03	< 2.00E-16	***
eparpip_30d	6.39E-03	1.12E-03	1.29E-08	***
epaedhc	-7.09E-02	2.70E-03	< 2.00E-16	***
epaedhd	-4.52E-01	8.42E-03	< 2.00E-16	***
ermtw_nm_per100f	-6.04E-03	1.53E-04	< 2.00E-16	***
epaed_copd_emphy_cb_18	5.32E-02	5.06E-03	< 2.00E-16	***
eparhesc	5.25E-02	1.37E-03	< 2.00E-16	***
leab10_15	1.82E-02	1.03E-03	< 2.00E-16	***
epar14_ormdpph_p30d_18	1.86E-01	3.74E-03	< 2.00E-16	***
epaedd18	1.66E-01	1.61E-03	< 2.00E-16	***
pm	-1.52E-01	2.55E-03	< 2.00E-16	***
oz	-7.63E-03	8.44E-04	< 2.00E-16	***
factor(sessv18)3	2.64E-02	1.45E-02		0.068509
factor(sessv18)4	-1.66E-02	1.09E-02		0.129725
factor(sessv18)5	-2.13E-01	2.53E-01		0.399295
factor(sessv18)6	2.20E-02	1.83E-02		0.228718
factor(preg13_17)3	-1.61E+00	7.13E-01	0.02362	*
factor(preg13_17)4	-4.53E-01	4.68E-02	< 2.00E-16	***
factor(preg13_17)5	-7.43E-01	7.55E-02	< 2.00E-16	***
factor(preg13_17)6	9.51E-01	3.80E-02	< 2.00E-16	***
factor(preg13_17)7	2.72E+00	4.48E-02	< 2.00E-16	***
factor(preg13_17)8	1.43E+00	3.72E-02	< 2.00E-16	***
factor(preg13_17)9	2.27E+00	4.05E-02	< 2.00E-16	***
factor(preg13_17)10	2.42E-01	3.43E-02	1.56E-12	***
factor(preg13_17)11	-4.54E-01	4.11E-02	< 2.00E-16	***
factor(preg13_17)12	1.60E-01	3.35E-02	1.80E-06	***
factor(preg13_17)13	-5.14E-02	3.55E-02		0.14758
factor(preg13_17)15	1.24E+00	1.44E-01	< 2.00E-16	***
factor(preg13_17)16	2.94E+00	1.07E-01	< 2.00E-16	***
factor(preg13_17)17	1.30E+00	1.26E-01	< 2.00E-16	***
factor(preg13_17)18	2.17E+00	1.09E-01	< 2.00E-16	***
factor(preg13_17)19	2.02E+00	3.20E-01	2.51E-10	***
factor(preg13_17)20	5.09E-01	1.52E-01	0.000778	***
factor(preg13_17)22	2.13E-01	3.27E-02	8.10E-11	***
factor(preg13_17)23	8.43E-02	3.40E-02	0.013207	*
factor(preg13_17)24	1.30E-01	3.19E-02	4.53E-05	***
factor(preg13_17)25	7.31E-02	3.26E-02	0.025077	*
factor(swi18)3	-6.84E-02	1.38E-02	7.13E-07	***
factor(swi18)4	-5.80E-02	1.06E-02	4.60E-08	***
factor(swi18)5	NA	NA	NA	
factor(swi18)6	-3.60E-02	1.71E-02	0.034885	*
factor(cddr08)2	-3.29E-01	2.11E-01		0.119874
factor(cddr08)3	-1.85E-01	2.09E-01		0.378377
factor(cddr08)4	1.22E-01	2.09E-01		0.560545
factor(cddr08)5	-6.41E-02	2.09E-01		0.759624
factor(cddr08)6	5.13E-02	2.09E-01		0.806432
factor(cddr08)7	-2.15E-02	2.10E-01		0.918323
factor(cddr08)8	-1.47E-01	2.10E-01		0.483807
factor(cddr08)9	2.24E-01	2.10E-01		0.28535
factor(cddr08)10	3.31E-01	2.11E-01		0.115775
factor(cddr08)11	2.79E-01	7.53E-02	0.00021	***
factor(hddr08)2	-1.55E-01	2.10E-01		0.459886
factor(hddr08)3	3.63E-01	2.09E-01		0.083152
factor(hddr08)4	2.58E-01	2.09E-01		0.217331
factor(hddr08)5	4.29E-01	2.09E-01	0.040431	*
factor(hddr08)6	1.39E-01	2.09E-01		0.507411
factor(hddr08)7	2.07E-01	2.09E-01		0.321464
factor(hddr08)8	2.97E-01	2.09E-01		0.155412
factor(hddr08)9	2.00E-01	2.09E-01		0.340209
factor(hddr08)10	7.82E-01	2.11E-01	0.000207	***
factor(hddr08)11	NA	NA	NA	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 4: Coefficients, Standard Error, and P -Value results from adjusted linear regression model.

	Estimate	Standard Error	P Value
(Intercept)	5.45E+00	1.26E-01	< 2.00E-16***
epaplp15_19	1.05E-02	4.48E-04	< 2.00E-16***
epci15_19	7.30E-06	2.76E-07	< 2.00E-16***
emaap15_19	8.11E-03	7.72E-04	< 2.00E-16***
epaedhbp	4.87E-02	2.51E-03	< 2.00E-16***
eparo	-6.43E-02	1.61E-03	< 2.00E-16***
eparpip_30d	1.15E-02	1.10E-03	< 2.00E-16***
epaedhc	-8.67E-02	2.65E-03	< 2.00E-16***
epaedhd	-4.26E-01	8.44E-03	< 2.00E-16***
ermtw_nm_per100f	-3.15E-03	1.09E-04	< 2.00E-16***
epaed_copd_emphy_cb_18	7.90E-02	4.99E-03	< 2.00E-16***
eparhesc	6.20E-02	1.35E-03	< 2.00E-16***
leab10_15	1.28E-02	1.05E-03	< 2.00E-16***
epar14_ormdpph_p30d_18	1.73E-01	3.61E-03	< 2.00E-16***
epaedd18	1.63E-01	1.62E-03	< 2.00E-16***
pm	-1.29E-01	2.41E-03	< 2.00E-16***
oz	-7.96E-03	8.40E-04	< 2.00E-16***
factor(preg13_17)3	-1.36E+00	7.44E-01	0.067166 .
factor(preg13_17)4	-4.79E-01	4.88E-02	< 2.00E-16***
factor(preg13_17)5	-7.85E-01	7.87E-02	< 2.00E-16***
factor(preg13_17)6	1.10E+00	3.94E-02	< 2.00E-16***
factor(preg13_17)7	3.03E+00	4.62E-02	< 2.00E-16***
factor(preg13_17)8	1.63E+00	3.86E-02	< 2.00E-16***
factor(preg13_17)9	2.57E+00	4.19E-02	< 2.00E-16***
factor(preg13_17)10	3.21E-01	3.57E-02	< 2.00E-16***
factor(preg13_17)11	-3.46E-01	4.24E-02	3.27E-16***
factor(preg13_17)12	2.57E-01	3.48E-02	1.58E-13***
factor(preg13_17)13	5.51E-02	3.67E-02	0.133255
factor(preg13_17)15	1.14E+00	1.50E-01	3.36E-14***
factor(preg13_17)16	2.86E+00	1.11E-01	< 2.00E-16***
factor(preg13_17)17	1.13E+00	1.31E-01	< 2.00E-16***
factor(preg13_17)18	2.02E+00	1.13E-01	< 2.00E-16***
factor(preg13_17)19	1.98E+00	3.34E-01	2.84E-09***
factor(preg13_17)20	5.66E-01	1.58E-01	0.000339***
factor(preg13_17)22	2.56E-01	3.41E-02	5.90E-14***
factor(preg13_17)23	1.05E-01	3.53E-02	0.002963**
factor(preg13_17)24	1.52E-01	3.32E-02	5.14E-06***
factor(preg13_17)25	7.70E-02	3.39E-02	0.023238*
factor(svl18)3	-1.34E-01	1.21E-02	< 2.00E-16***
factor(svl18)4	-1.06E-01	9.99E-03	< 2.00E-16***
factor(svl18)5	-1.93E-01	2.63E-01	0.46324
factor(svl18)6	-1.67E-01	1.47E-02	< 2.00E-16***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5: Coefficients, Standard Error, and P-Value results from revised adjusted linear regression model.

	Estimate	Standard Error	P Value
(Intercept)	8.59E+00	1.34E-01	< 2.00E-16***
epaplp15_19	7.20E-03	4.22E-04	< 2.00E-16***
epci15_19	4.65E-06	2.57E-07	< 2.00E-16***
emaap15_19	8.01E-03	7.26E-04	< 2.00E-16***
epaedhbp	-1.51E-01	3.18E-03	< 2.00E-16***
eparo	-7.00E-02	1.58E-03	< 2.00E-16***
eparpip_30d	1.43E-02	1.05E-03	< 2.00E-16***
epaedhc	6.75E-02	2.97E-03	< 2.00E-16***
epaedhd	-1.24E-01	8.77E-03	< 2.00E-16***
ermtw_nm_per100f	-2.69E-03	1.03E-04	< 2.00E-16***
epaed_copd_emphy_cb_18	1.16E-01	4.80E-03	< 2.00E-16***
eparhesc	4.53E-02	1.35E-03	< 2.00E-16***
leab10_15	2.14E-02	9.93E-04	< 2.00E-16***
epar14_ormdpph_p30d_18	2.10E-01	3.72E-03	< 2.00E-16***
epaedd18	1.55E-01	1.59E-03	< 2.00E-16***
pm	-1.20E-01	2.31E-03	< 2.00E-16***
oz	-2.06E-03	8.12E-04	0.0111*
blkp	1.98E+00	7.11E-02	< 2.00E-16***
whip	-3.90E+00	6.69E-02	< 2.00E-16***
hisp	-3.87E+00	6.75E-02	< 2.00E-16***
asip	-5.66E+00	9.08E-02	< 2.00E-16***
factor(svl18)3	-2.94E-02	1.15E-02	0.0105*
factor(svl18)4	-4.91E-02	9.48E-03	2.18E-07***
factor(svl18)5	-5.27E-02	2.49E-01	0.8325
factor(svl18)6	-2.80E-02	1.38E-02	0.0426*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Discussion

The findings of this study are consistent with previous studies analyzing the associations of environmental quality and adult asthma prevalence [6]. We found that there are higher rates of asthma in neighborhoods that have high rates of exposure to contaminants such as particulate matter. These findings substantiate previous public health research on outdoor air pollution and

increased rates of asthma [7]. Furthermore, we found that there is a positive correlation between individual factors and increased rates of asthma. For example, we found that there is a positive correlation between estimated percent of adults reporting to have asthma and neighborhood makeup in terms of percent black people living in the neighborhood. This finding is consistent with data published by the Asthma and Allergy Foundation of America [1].

For the purposes of this study we utilized an all variable linear regression model to identify variables and factors that have significance. We used these identified significant variables and factors in an adjusted variable linear regression. Over 50% of the variables used in this study held significance. The variables that hold significance can be found in Table 2. Furthermore, two of five factors used in this study held significance. The variables that are not significant may be attributable to the reason that this study is focused on evaluating neighborhood and environmental exposures and asthma. Therefore, we are not accounting for individual-level risk factors, which may have a large effect on asthma rates that are not captured by our model. In our revised adjusted linear regression model we found that for a one unit increase in estimated percent of adults reporting to have 14 or more days of poor physical health in the past 30 days in 2018 there is a 0.210 unit increase in estimated percent of adults reporting to have asthma in 2018. This finding is consistent with our expectations of poor health being associated with increased asthma prevalence. We also found in our model that for a one unit increase in percent Black people in a neighborhood there is a 1.98 unit increase in estimated percent of adults reporting to have asthma in 2018. This finding and direction of this coefficient is consistent with our expectations as previous studies have substantiated that Black people have 1.25 times the asthma prevalence when compared to the US general population [14]. For a one unit increase in particulate matter and ozone there is a decrease of 0.120 and 0.00206 units

respectively in estimated percent of adults reporting to have asthma in 2018. The direction of these findings is not consistent with our expectations as previous research has substantiated that long term exposure to ozone and particulate matter is associated with respiratory diseases [10].

The findings of our study are important because they serve to substantiate previous public health research on different parts of the neighborhood environment, with a focus on air pollution as it correlates to increased asthma incidence on a neighborhood scale. Moreover, the results of this study hold statistical significance, which can be used to inform data driven decisions regarding the enactment of policy with an emphasis on environmental justice.

The strengths of this analysis reside in applying analytical approaches on neighborhood levels of asthma rates while evaluating the characteristics of these neighborhoods, in order to identify vulnerable neighborhoods and the characteristics that are associated with high prevalence of asthma. Another strength of this study is that we created a dataset at the census tract level, by collecting data from multitudes of sources. Literature and previous studies infrequently study the associations between environmental quality, behavioral factors, socioeconomic status, educational attainment, neighborhood characteristics, modifiable risk factors, and comorbidities to asthma prevalence. The weaknesses of this study are that this study is focused on evaluating neighborhood and environmental exposures and asthma, we are not accounting for individual-level risk factors, which may have a large effect on asthma rates that are not captured by our model. This may explain the R-squared values and the overall fit of our model, along with insignificant variables in our all variable linear regression model. This study fills a research gap in terms of analyzing neighborhood and environmental predictors of asthma in the U.S.A.

Conclusion

The use of advanced statistical and spatial methods of analysis on asthma prevalence as it correlates on a neighborhood scale to environmental quality, behavioral factors, socioeconomic status, educational attainment, neighborhood characteristics, modifiable risk factors, and comorbidities is an area of sparse research. By cultivating a greater understanding of how these factors interact simultaneously the strategic optimization of neighborhood level interventions can be implemented for asthma prevention. Comprehensive datasets at the census level from multiple sources maximizes the impact of interventions at the community scale. Coupling these comprehensive datasets with advanced statistical and spatial analysis while emphasizing environmental justice will be needed to improve asthma prevalence rates and adapt to climate change.

References

1. AAFA [Internet]. Asthma and Allergy Foundation of America. [cited 2022May2]. Available from: https://www.aafa.org/asthma-facts/
2. CDC - BRFSS [Internet]. Centers for Disease Control and Prevention. Centers for Disease Control and Prevention; 2022 [cited 2022May2]. Available from: https://www.cdc.gov/brfss/index.html
3. D'Amato G, Cecchi L, D'Amato M, Annesi-Maesano I. Climate change and respiratory diseases [Internet]. European Respiratory Society. European Respiratory Society; 2014 [cited 2022May2]. Available from: https://err.ersjournals.com/content/23/132/161.full
4. Does outdoor air pollution induce new cases of ... - Wiley Online Library [Internet]. [cited 2022May3]. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1440-1843.2012.02195.x
5. EPA. Environmental Protection Agency; [cited 2022May2]. Available from: https://www.epa.gov/newsreleases/epa-report-shows-disproportionate-impacts-climate-change-socially-vulnerable
6. Gray CL, Lobdell DT, Rappazzo KM, Jian Y, Jagai JS, Messer LC, et al. Associations between environmental quality and adult asthma prevalence in medical claims data [Internet]. Environmental Research. Academic Press; 2018 [cited 2022May2]. Available from: https://www.sciencedirect.com/science/article/abs/pii/S001393511830327X?via%3Dihub
7. Guarnieri M, Balmes JR. Outdoor Air Pollution and asthma [Internet]. Lancet (London,

England). U.S. National Library of Medicine; 2014 [cited 2022May2]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4465283/

8. Hurbain P, Liu Y, Strickland MJ, Li D. A cross-sectional analysis of associations between environmental indices and asthma in U.S. counties from 2003 to 2012 [Internet]. Nature News. Nature Publishing Group; 2021 [cited 2022May2]. Available from: https://www.nature.com/articles/s41370-021-00326-4
9. Indoor Air Quality [Internet]. Indoor Air Quality: 4. Are certain people more vulnerable than others to indoor air pollution? [cited 2022May2]. Available from: https://ec.europa.eu/health/scientific_committees/opinions_layman/en/indoor-air-pollution/l-2/4-vulnerable-groups.htm#:~:text=Elderly%20people%20may%20be%20particularly,of%20the%20eyes%20and%20nose
10. Lim C. Long-Term Exposure to Ozone and Cause-Specific Mortality Risk in the United States [Internet]. American Journal of respiratory and Critical Care Medicine. 2019 [cited 2022May2]. Available from: https://www.atsjournals.org/doi/full/10.1164/rccm.201806-1161OC
11. Oland AA, Booster GD, Bender BG. Psychological and lifestyle risk factors for asthma exacerbations and morbidity in children [Internet]. The World Allergy Organization journal. BioMed Central; 2017 [cited 2022May2]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5644196/#:~:text=A%20range%20of%20risk%20factors,physical%20inactivity%2C%20and%20unhealthy%20diets
12. Porter D, Nash E, McCullough M. [Internet]. PolicyMap. 2021 [cited 2022May2]. Available from: https://www.policymap.com/

13. Staff. AZ asthma advice: 15 worst U.S. cities for Asthma - Phoenix: AAPA [Internet]. AZ Asthma Advice |15 Worst U.S. Cities For Asthma - Phoenix | AAPA. 2019 [cited 2022May2]. Available from: https://allergyarizona.net/high-desert-asthma-and-allergies-advice-from-asthma-specialists/

14. Sullivan K, Thakur N. Structural and social determinants of health in asthma in developed economies: A scoping review of literature published between 2014 and 2019 [Internet]. Current allergy and asthma reports. Springer US; 2020 [cited 2022May2]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7005090/#CR1

15. Szentpetery SS, Gruzieva O, Forno E, Han Y-Y, Bergström A, Kull I, et al. Combined effects of multiple risk factors on asthma in school-aged children [Internet]. Respiratory medicine. U.S. National Library of Medicine; 2017 [cited 2022May2]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5728683/#:~:text=Indeed%2C%20current%20evidence%20suggests%20causal,and%20outdoor%20pollutants%20(4)