UNIVERSITY OF CALIFORNIA,
IRVINE

Association between PFOA/PFOS Drinking Water Exposures and Asthma ED Visits

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Environmental Health Sciences

by

Thomas Matthew Overton

Thesis Committee:
Professor Veronica Vieira, Chair
Professor Dean Baker
Professor Ulrike Luderer

2022

DEDICATION

To

my family and friends

in recognition of their support

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express gratitude and appreciation to my committee chair, Professor Veronica Vieira, for her thoughtful guidance in both conducting the research and writing this thesis. I would like to thank my committee members, Professor Dean Baker and Professor Ulrike Luderer, for their support of this research project as well as for their insightful comments and suggestions. I would also like to thank Dr. Alya Khan, residency program director, for her ongoing advice and support throughout the Occupational Medicine residency program.

ABSTRACT OF THE THESIS

Association between PFOA/PFOS Drinking Water Exposures and Asthma ED Visits

by

Thomas Matthew Overton

Master of Science in Environmental Health Sciences

University of California, Irvine, 2022

Professor Veronica Vieira, Chair

Perfluoro-n-octanoic acid (PFOA) and Perfluoro-1-octanesulfonate (PFOS) are ubiquitous and highly stable toxins in our environment. Animal and epidemiological research studies have begun to shed light on the toxicological profile and health effects of PFOA/S, including their impacts on respiratory function. Asthma is one of the health effects that has been linked to PFOA/S exposure. In this study, we explored the association between PFOA/S in drinking water and asthma emergency department (ED) visits per 10,000 in Orange (OC) and LA counties (LAC), California. These studies were conducted at the population level using data sets containing drinking water concentrations of PFOA/S and number of asthma ED visits within the ZIP Codes of the 2 counties. PFOA/S concentrations in drinking water for LAC and OC water districts from 2013-2015 was obtained from the Third Unregulated Contaminant Monitoring Rule dataset. We conducted geographical analysis using ArcGIS by quantitatively mapping PFOA/S drinking water levels and asthma ED visits by ZIP Code. We performed statistical analysis using linear as well as non-linear regression methods to determine if PFOA/S drinking water levels were associated with asthma ED. These statistical analyses controlled for various potential confounders of this relationship — race,

education, household income, air quality (diesel, $PM_{2.5}$). Mapping and statistical analysis did not reveal a direct relationship between PFOA/S drinking water levels and rates of asthma ED visits.

Chapter 1: Introduction

Per-/Poly - fluoroalkyl substances (PFAS) are increasingly recognized contaminants in drinking water. This presents a public health problem as growing evidence has linked PFAS to human health problems, including asthma. While the research literature provides supporting evidence of the association between PFAS and asthma, there is little insight into the effects of PFAS on asthma severity and exacerbations.

Asthma and Asthma Exacerbations

Asthma is one of the most common respiratory illnesses in the United States of America. The National Health Institute Survey estimates that 25 million Americans have asthma. This estimate includes 8% of adults and 7% of children. It is well-known that asthma is a chronic disease marked by acute exacerbations. These acute episodes, some of which can be life-threatening, disproportionately affect children (Figure 1). In 2019, 44.3% of children with asthma were reported to have had asthma attacks (NHIS, 2019). Common causes of asthma exacerbation include reactions to allergens (such as pollen, pets, mold, and dust mites), upper respiratory infections, and tobacco smoke, amongst other factors.

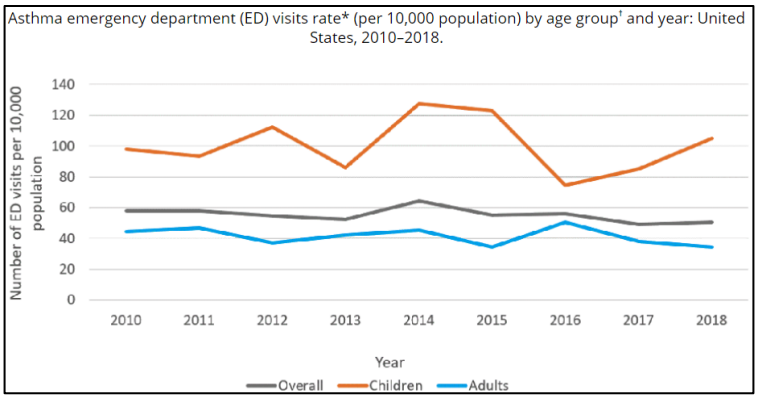


Figure 1. Asthma emergency department visit rate (per 10,000 population) by age group and year: United States, 2010 - 2018. Children (orange line), adults (blue line), and overall population (Black line). (NHIS, 2019)

Contributing Factors to Asthma Exacerbations

Sociodemographics

In the United States and worldwide, socioeconomic status is a major determinant of health, including for asthma. In California from 2011–2013, historically redlined census tracts had 2.4 times higher rates of emergency department visits per year due to asthma compared to tracts at the lowest risk level of discrimination (Nardone et al., 2021). Communities with lower socioeconomic status, including disadvantaged minorities, often have lower continuity of care with primary care physicians and consequently have higher rates of ED visits for chronic conditions. Along these lines, increased asthma-related emergency department (ED) visits are associated with fewer asthma-related primary care visits leading up to the ED visit (Smith et al., 2007). In 2000, the odds that Connecticut children with 3 or more asthma-related primary care visits and a filled inhaled corticosteroid (ICS) had an asthma ED visit were one-fifth (OR = 0.20, 95% CI (0.06, 0.65)) those of children with fewer asthma-related PCP visits or no filled ICS. Furthermore, this study showed that fewer than 17% of children with asthma-related ED visits filled any prescription for

controller medications in the 3 months before these visits (Smith et al., 2007). Increased asthma ED visits are not only the result of not receiving proper treatment for their asthma but also not receiving education from their doctor about their medical condition. No education about asthma and uncontrolled asthma are major factors leading to frequent asthma-related ED visits (Hamdan et al., 2012).

The burden of asthma prevalence in the United States falls disproportionately on Black and American Indian/Alaska Native people. These groups have the highest current asthma prevalence rates compared to other races and ethnicities. According to the American Lung Association, the prevalence of asthma among Blacks and American Indian/Alaska Native people in 2018 was 10.9% and 12.0%, respectively. This is in comparison to much lower rates in Whites (7.7%), Hispanics (6.4%), and Asians (4.0%). In addition to asthma prevalence, minority groups, particularly African-Americans, also face disparities in asthma exacerbations (Guillbert et al., 2019). Asthma exacerbations are the focus of this study.

Air Quality

The respiratory tracts of individuals who have asthma is particularly sensitive to air pollutants and irritants that may be inhaled. Air quality is a measure of the level of pollutants in the air. These pollutants include carbon monoxide, lead, nitrogen oxides, sulfur dioxide, ozone, particulate matter, and diesel exhaust. In many cases these air pollutants are emitted as a result of combustion from vehicles, stoves, heaters, power plants, oil refineries, industrial facilities, and factories. The two confounding air quality variables that were included in the final analysis were $PM_{2.5}$ and diesel exhaust, which is a

major source of PM2.5. Particulate matter 2.5 ($PM_{2.5}$) or fine particulate matter refers to tiny particles or droplets in the air that are two and one half microns or less in aerodynamic diameter allowing them to enter into the smaller airways of the lungs. Significant associations have been found between ED visits for asthma in children and fine particulate matter (Norris et al., 1999). Diesel exhaust is the gaseous exhaust produced by a diesel type of internal combustion engine. Los Angeles County low-income communities of color experience uneven asthma ED visits, partly attributed to inequitable exposure to diesel particulate matter (Douglas et al., 2019).

Per-/Poly - fluoroalkyl substances

Per-/Poly - fluoroalkyl substances are a group of man-made chemicals that are ubiquitous in the environment. The water and lipid resistant properties of PFAS have resulted in their widespread use in industrial surfactants, firefighting foams, and textile treatments (Chang et al., 2016). As a result of their strong carbon-fluorine (C-F) backbone (Figure 2), PFAS are environmentally persistent chemicals (Savvaides et al., 2021). The C-F backbone makes PFAS resistant to thermal and chemical stress in the environment and within biological organisms. Consequently, the biological half-life of PFAS in humans is on the order of years. PFAS have been associated with a number of health effects in humans including reproductive and developmental, liver, kidney, and immunological effects. The most commonly studied PFAS are Perfluoro-n-octanoic acid (PFOA) and Perfluoro-1-octanesulfonate (PFOS) (Figure 2).

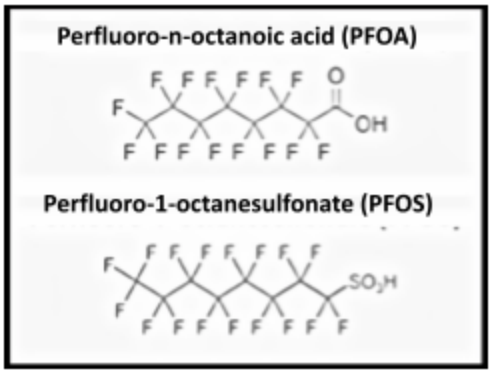


Figure 2. PFOA and PFOS. Molecular structure of PFOA and PFOS.

PFOA/S and Drinking Water

PFAS are widespread drinking water contaminants because they are highly persistent, mobile in groundwater, and bioaccumulate. PFAS were found in the drinking water of more than 16 million Americans in 33 states (Hu et al., 2016). Seven states have developed their own water guideline levels for PFOA and/or PFOS ranging from 13 to 1000 ng/L, compared to the EPA's health advisory (HAs) of 70 ng/L for both compounds individually or combined (Cordner et al., 2019). Public water systems (PWSs) are not required to routinely test for PFAS or to treat water exceeding EPA HAs, and so no complete assessment of the prevalence of PFAS in U.S. drinking water exists (Cordner et al., 2019). Furthermore, California, the location of this study, uses a large amount of recycled water, a volume that totaled 686,000 acre-feet in 2019. The large use of recycled water combined with the thermal/chemical stability of PFAS has the potential to promote large accumulation of PFAS overtime.

PFOA/S and Asthma

Molecular Mechanism of Asthma

The immune system plays an essential role in defending the body against pathogens that enter the body. Epidemiological studies have identified associations between PFOA/A exposure and immune responses and asthma. Moreover, experimental toxicological evidence also suggests that PFOA affects the immune response. When an inappropriate immune response occurs in response to an antigen or allergen, the process is referred to as a hypersensitivity reaction.

There are four types of hypersensitivity reactions -- Type I is mediated by IgE antibodies, Type II is a cytotoxic reaction mediated by IgG or IgM antibodies, Type III is mediated by immune complexes, and Type IV is a delayed reaction mediated by cellular response. Asthma is a type 1 hypersensitivity reaction. In a type 1 hypersensitivity reaction, antigen presenting cells, such as dendritic cells, bring the allergen to T2 helper cells (T_H2). This triggers T_H2 cells to release cytokines that activate B cells. Activated B cells produce and release IgE antibodies specific to the allergen. IgE binds to the allergen forming a complex. This IgE-allergen complex then binds to receptors on mast cells causing mast cells to degranulate and release the molecules that produce the allergic reaction.

<u>Associations between asthma risk factors and PFAS exposure</u>

The apparent associations between PFAS exposure and asthma exacerbations (cite literature) may be confounded by associations between PFAS exposure and known asthma risk factors. Individual studies have shown that human biomonitoring biomarker

concentrations of PFAS generally increase with income. A 2018 meta-analysis performed by Buekers et al. showed that a higher income is associated with a higher internal exposure to PFOA/S, as well as other PFAS subtypes. As income doubled, internal exposure to PFAS increased on average by 10%–14% (Buekers et al., 2018). Descriptors for education (<college, x years of study, etc.) differed too widely between studies to perform a meta-analysis.

Racial/ethnic burdens differ between PFOS and PFOA. White women had higher concentrations of PFOA compared with Chinese in Oakland and Blacks in Pittsburgh, while Black women had higher concentrations of PFOS compared to White women in Southeast Michigan and Boston (Park et al., 2019). PFAS can be found in $PM_{2.5}$ with PFOS serving as the predominant PFAS subtype in $PM_{2.5}$ collected from ambient air (Zhou et al., 2021).

Research Objectives

This project explores the two most commonly studied PFAS, Perfluoro-n-octanoic acid (PFOA) and Perfluorooctanesulfonic acid (PFOS), and their level in drinking water in relation to the number of asthma-related emergency department (ED) visits per 100,000 in Orange and Los Angeles counties by ZIP Code. There is a conceivable association between PFAS exposure and asthma exacerbation. Proof of such an association would merit exploration of policy changes to mitigate PFAS prevalence. This study aims to extend the current research literature and to inform future studies. The primary hypothesis is that higher levels of PFAS in drinking water leads to a greater rate of asthma ED visits. Moreover, there are three main objectives of the study:

1. Determine PFAS drinking water concentrations by ZIP Code using known concentrations in water districts.
2. Use ArcGIS to map PFAS drinking water concentrations by ZIP Code.
3. Determine if there is a statistical association between PFAS concentrations in drinking water and rates of asthma ED visits, while controlling at the ZIP code level for potential confounding factors.

Chapter 2: Literature Review

Animal Studies

Immune system involvement of PFAS

Early toxicological studies have provided evidence to support the hypothesis that PFOA/S affects the immune system. Studies have demonstrated that PFOS can activate pathways of the innate immune system (Qazi et al., 2009). Alternatively, other studies have shown that PFOS may suppress acquired immunity (DeWitt et al. 2008; Peden-Adams et al. 2008; Dong et al. 2009) and PFOS has been shown to be immunosuppressive to mice (Peden-Adams et al., 2008). These research studies assessed immune system changes for PFOS exposure by mostly evaluating the weight of immune organs and the numbers of lymphocytes.

In 2011 Dong et al. performed studies in a mouse model that suggest PFOS exposure shifts the host immune response from a balanced TH1/TH2 state toward a dominant TH2 state, a hallmark of atopic disease (Dong et al., 2011). In these experiments C57BL/6 mice were exposed to PFOS daily via gavage for 60 days. One day after the final exposure, the production of TH1-type cytokines (IL-2 and IFN-c) and TH2-type cytokines (IL-4) by isolated splenocytes were assessed via ELISA or ELISPOT. The secretion of the TH-2 cytokine IL-4 was increased in a dose-dependent manner whereas PFOS exposure decreased the formation of the type 1 cytokines IL-2 and IFN-c. Serum levels of immunoglobulin (Ig) were measured in a similar manner. Researchers observed significantly increased serum (sheep red blood cells (SRBC)-specific) IgE, IgG, and IgG1

levels. Alternatively, significantly decreased (SRBC-specific) IgM synthesis was observed with decreased PFOS exposure in a dose-related manner.

PFAS and Asthma

Studies using animal models have provided data suggesting that PFOA/PFOS exposure has the potential to exacerbate Type 1 or IgE-mediated hypersensitivity responses, such as asthma. These studies have demonstrated increased IgE levels and airway hypersensitivity in asthma animal models. While there have been several species of animals used to model asthma, by far the most common model has been mice, particularly BALB/c (Aun et al., 2017). In order to mimic human pathophysiology, animal models of asthma usually have two phases: sensitization and challenge with allergen. The traditionally used allergen has been the main protein found in egg white, ovalbumin, which causes an intense inflammatory response. To improve the relevance to human disease, ovalbumin is being replaced by aeroallergens, such as house dust mites, which are more common causes of asthma exacerbations. The outcomes measured include serum-specific IgE antibodies, airway hyperresponsiveness, inflammation and remodeling. In many of the animal studies airway hypersensitivity is assessed by methacholine challenge. Methacholine is a chemical that causes bronchoconstriction. For airways that are highly sensitive, such as the case for asthmatics, this bronchoconstriction occurs to a greater extent.

In 2007 Fairley et al. observed that PFOA increased serum levels of IgE and enhanced hypersensitivity responses to environmental allergens in a murine model (Fairley et al., 2007).

The IgE response was measured in BALB/c mice after coexposure to the allergen ovalbumin and increasing concentrations of PFOA over a 14-day period. The mice received an intraperitoneal injection (7.5 mg) of ovalbumin. Mice were then separated and given an injection into the dorsal surface of each ear with different concentrations of PFOA. In the presence of ovalbumin, increasing concentrations of PFOA resulted in significantly increased levels of both total IgE and ovalbumin-specific IgE. These mice were also assessed for airway hyperreactivity. Mice were placed in a plethysmograph, an instrument that can be used for measuring airway volumes. PenH values, a dimensionless index used to evaluate changes in the shape of the airflow pattern entering and leaving a whole-body flow plethysmograph as an animal breathes, were measured for a 5-hour time period. In the presence of ovalbumin, increasing concentrations of PFOA resulted in increased airway hyperreactivity.

Zeng et al. published a study that demonstrated PFOA increases airway hypersensitivity and downregulates glucocorticoid receptor expression in asthmatic mice (Zeng et al., 2021). Currently, inhaled glucocorticoids are the first-line treatment for asthma and their binding to glucocorticoid receptors is essential to their anti-inflammatory effect. BALB/c mice were administered ovalbumin (OVA) and PFOA was then administered intratracheally for seven days. Increasing concentrations of PFOA resulted in increased airway hypersensitivity to methacholine. Furthermore, increased concentrations of PFOA resulted in decreased glucocorticoid receptor mRNA and protein in the lungs of mice. This was an interesting finding, as patients with severe asthma show reduced expression of glucocorticoid receptors (Chang et al., 2012). In a previous study, PFOA was shown to

inhibit 11β-hydroxysteroid dehydrogenase type 1, which catalyzes reactions that increase the expression of glucocorticoid receptor (Ye et al., 2021; Chapman et al., 2013).

Another recent study explores a potential mechanistic explanation for the role of PFAS in increased airway hyperreactivity. Wang et al., showed that PFOS abolishes the ability of lipopolysaccharide (LPS), a gram-negative bacterial membrane endotoxin, to protect against asthma development (Wang et al., 2021). Authors used a BALB/c mouse model for asthma which utilized house dust mite (HDM) intratracheally as the allergen. LPS alone was able to reduce IgE and cytokine production, immune cell recruitment, and methacholine sensitivity of mice in response to HDM. The addition of PFOS reduced the ability of LPS to protect against IgE and cytokine production, immune cell recruitment, and methacholine sensitivity in response to HDM. LPS is one of the most potent microbe-associated molecular pattern molecules and TLR4 agonists. LPS is present at high levels in air and dust, and LPS exposure negatively correlates with the risk of developing asthma (Ege et al., 2011; Mutius et al., 2010; Riedler et al., 2001). Regular exposure to LPS up-regulates the expression of a negative regulator (protein A20) of airway epithelial cell inflammation. Thus, LPS reduces the production of cytokines/chemokines in response to allergens and in turn suppresses dendritic cell recruitment and the T helper 2 response that results in asthma (Schuijs et al., 2015; Lee et al., 2000). In addition to direct airway interactions, airway inflammation is also affected indirectly by LPS. The intestinal microbiota modulates allergic asthma, and a low abundance of intestinal LPS is associated with a high risk of airway inflammation (Arrieta et al., 2015). The route of ingestion of PFOS through contaminated water may influence airway inflammation through its effects on intestinal LPS. Although Wang et al., showed PFOS + LPS + HDM to increase airway

hyperreactivity compared to LPS + HDM, PFOS alone reduced airway hyperreactivity in response to HDM alone.

Gestational exposure to PFAS

Ryu et al demonstrated that PFOA/S increased airway resistance in animal models (Ryu et al., 2014). In these experiments, pregnant dams were fed PFOA- or PFOS-contaminated diet ad libitum (~4–6 g/day) beginning on gestation day (GD) 2 through the pregnancy and lactation periods. Offspring at 8–9 wk of age were sensitized by intraperitoneal injection of ovalbuin (2 μg). To measure lung mechanics, mice were anesthetized and tracheotomized with a catheter connected to a flexiVent small animal ventilator (Scireq Montreal, PQ, Canada). Lung mechanics were measured using a mechanical ventilator. Airway mechanics were assessed after inhalation of increasing concentrations of methacholine. Methacholine increased airway resistance in a dose-dependent manner as expected and PFOA/S alone was able to increase the sensitivity of mice to methacholine-induced airway resistance. Furthermore, the addition of PFOA significantly increased the sensitivity of methacholine-induced airway resistance above that of the ovalbumin exposed mice alone. The mechanisms by which prenatal exposures cause asthma in offspring would likely differ compared to direct exposure.

Epidemiological Studies

Epidemiological studies have provided data suggesting that PFOA/S exposure has the potential to exacerbate Type 1 or IgE-mediated hypersensitivity responses, such as asthma in humans.

GBCA-based Epidemiology Studies

The Genetic and Biomarkers study for Childhood Asthma (GBCA) was conducted in Taiwan from 2009 to 2010. Study participants were recruited from two hospitals in Taipei City of Northern Taiwan. The study subjects were children aged 10 - 15 years old who were diagnosed with asthma within the previous year. This resulted in 231 asthmatic subjects. The control population was recruited from seven public schools in Taipei City of Northern Taiwan. Age-matched controls with no personal or family history of asthma were chosen. This totaled 225 non-asthmatic controls.

A study by Dong et al. (2013) used data from the GBCA. Logistic regression models adjusting for age, sex, BMI, parental education, ETS exposure, and month of survey showed increased odds ratios (OR) for asthma with PFOA/S serum concentrations. The OR for PFOA was 4.05 (95% CI: 2.21, 7.42) and PFOS was 2.63 (95% CI: 1:48, 4.69), respectively comparing the highest quartile to the lowest quartile serum concentrations. Also higher asthma severity scores among the children with asthma were associated with increasing serum PFOA and PFOS (Dong et al., 2013). Among the asthmatic children, asthma severity scores were determined using a 13-item questionnaire to evaluate for frequency of current asthma symptoms, use of systemic corticosteroids, use of medications besides systemic corticosteroids, and history of hospitalizations and intubations. The asthma severity scoring system ranged from 0 to 28, with higher scores representing more severe asthma. Asthmatic children with increased serum PFOS had significantly higher asthma severity scores compared to asthmatic children with lower serum PFOS. Asthma severity scores divided up by PFOS quartiles were quartile 1 - 3.33 (2.36, 4.31); quartile 2 - 4.18 (3.19,

5.17); quartile - 3 4.49 (3.52, 5.45); quartile 4 4.57 - (3.61, 5.54)]. Although not significant, asthma severity scores showed a positive trend with increasing PFOA quartiles. Dong et al.'s publication also showed increasing serum IgE and eosinophil levels (asthma related-biomarkers) with increasing serum PFOA and PFOS (Dong et al., 2013).

In a study by Qin et al., results of spirometry testing on a subset of the GBCA population suggested that increasing serum concentration of PFOA and PFOS is associated with decreased lung function in children with asthma but not in children without asthma (Qin et al., 2017). Lung function parameters assessed were 1) forced expiratory volume (FEV), which measures the quantity of air a person can exhale during a forced breath, 2) forced vital capacity (FVC), which measures the quantity of air a person can forcibly exhale from their lungs after taking the deepest breath possible, 3) forced expiratory flow 25–75% (FEF25-75), which measures the quantity of air a person can exhale from the point at which 25% of the FVC has been exhaled to the point at which 75% of the FVC has been exhaled. Linear regression modeling adjusted for age, sex, BMI, parental education level, exercise, ETS exposure, and month of survey showed serum PFOA/S levels were significantly negatively associated with three pulmonary function measurements in children with asthma. The adjusted coefficients between lung function and PFOA/Ss exposure included the following FEV: −0.104 (95%CI: −0.193 to −0.015) for PFOA and −0.061 (95%CI: −0.101 to −0.021) for PFOS; FVC: −0.067 (95%CI: −0.167 to 0.032) for PFOA and −0.055 (95%CI: −0.100 to −0.010) for PFOS; FEF25-75: −0.223 (95%CI: −0.400 to −0.045) for PFOA and −0.045 (95%CI: −0.127 to 0.037) for PFOS. Coefficients represent the mean change in

pulmonary function measurements for each 1 ln-(ng/mL) increase in serum PFOA/S concentration.

Zhu et al. generated data from the GBCA cross-sectional study. This study looks at the same study subjects and measures as the Dong et al. study, just limited to boys. The results of the Zhu et al. study suggests that among males with increasing serum PFOA and PFOS there is an association with increasing odds of asthma (Zhu et al., 2016). Models were adjusted for age, BMI, parental education, ETS exposure, parental asthma, and month of survey. Among males, adjusted odds ratios for asthma (4th quartile vs 1st quartile of PFOA/S exposure) were 4.24 (95% CI: 1.91, 9.42) for PFOA and 4.38 (95% CI: 2.02, 9.50) PFOS. Just as in previous animal models (Dong et al., 2011) Zhue et al. provided epidemiologic evidence for the ability of PFAS to shift the host immune response from a balanced T_H1/T_H2 state toward a dominant T_H2 state, an immunological state that increases the likelihood of an asthma exacerbation. Serum concentrations of PFAS and levels of T_H1 (interferon-γ, IL-2) and T_H2 (IL-4 and IL-5) cytokines were measured in children aged 10-15 years-old who were diagnosed with asthma within the previous year. Serum PFAS, including PFOA/S, were associated positively with T_H2 cytokines and inversely with T_H1 cytokines among male asthmatics, but not females. This finding makes sense as prepuberal boys are more likely to have diagnosed asthma and be hospitalized from asthma symptoms (Fuseini and Newcomb, 2017). Furthermore, increasing serum concentrations of PFOA/S were associated with increasing serum IgE levels. Among females, significant increasing odds of asthma were found only with PFOA, and not PFOS. Adjusted odds ratios for asthma (4th quartile vs 1st quartile of PFOA/S exposure) for females was 3.68 (95% CI: 1.43, 9.48) for PFOA. Among

females, there were no significant associations between PFOA/S and T_H2 cytokines and T_H1 cytokines identified.

NHANES-based Epidemiology Studies

The National Health and Nutrition Examination Survey (NHANES) is conducted by the U.S. Centers for Disease Control and Prevention (CDC). NHANES assesses health and nutrition in a nationally representative sample of the US population. It consists of interviews and physical examinations of approximately 5,000 persons each year. These data are released every 2 years. With regards to PFOA/S and asthma, NHANES measures serum PFOA/S concentrations using analytical chemistry methods and asthma is assessed by parent-reported, doctor-diagnosed, asthma using a standardized questionnaire.

A cross-sectional study by Humblet et al. (2014) used data from the NHANES (1999–2000 and 2003–2008) to evaluate the association of PFAS with reported asthma. This study used data from 4 different NHANES: 1999–2000, 2003–2004, 2005–2006, and 2007–2008. Multivariable models adjusting for sex, smoking, age, race/ethnicity, survey cycle, poverty income ratio, and health insurance evaluated the effect of a doubling of PFAS concentrations on asthma. These data showed that increasing serum PFOA was associated with an increased odds 1.18 (95% CI: 1.01, 1.39) of parent-report of ever receiving a diagnosis of asthma among children 12–19 years of age (Humblet et al., 2014). On the other hand, there was an inverse relationship observed for PFOS and ever asthma; OR = 0.88; 95% CI: 0.74, 1.04 (Humblet et al., 2014).

In a subset of this population (NHANES 2005–2006), Stein et al. observed no adverse association between PFOA/S exposure and asthma. Models were adjusted for age, gender, race, weight status, and serum cotinine. Adjusted odds ratios for asthma (75th percentile vs 25th percentile of PFOA/S exposure) were calculated. These data showed elevated ORs for PFOA of 1.28 (95% CI: 0.81, 2.04) and PFOS of 1.20 (95% CI: 0.88, 1.63) with asthma, but with wide confidence intervals (Stein et al., 2016).

From a more recent NHANES (2013–2014) dataset, Jackson-Browne et al. (2020) found serum PFAS concentrations were weakly associated with increased asthma prevalence in US children. Unlike the two previous NHANES analyses (Humblet et al., 2014 and Stein et al., 2016) which focus on 12–19 year old subjects, this NHANES analysis examined a younger population of children aged 3–11 years. Regression models were adjusted for sex, age, race/ethnicity, serum cotinine, and poverty to income ratio. Odds ratios for the outcome of asthma were calculated per standard deviation increase in ln-transformed serum PFOA/S concentrations. This analysis showed increased serum PFOA (1.1; 95% Cl: 0.8, 1.4) and PFOS (1.2; 95% Cl: 0.8, 1.7) to be weakly associated with an increased odds of asthma. This study also examined effect measure modification of age, sex, and race. This analysis provided evidence to support the ability of age to modify the associations between serum PFOS, but not PFOA, concentrations and odds of asthma. Sex and race did not modify these associations (Jackson-Browne et al., 2020).

Chapter 3: Methods

Study Areas

Los Angeles County, CA and Orange County, CA were selected as the areas of study for the analysis and estimation. The study areas are shown in (Figure 3). LA County and Orange County are contiguous counties in the southern region of California. Los Angeles County is the most populous county of both the United States and the state of California with a population of over ten million. LA County covers over 4000 sq miles of land territory, and is broken into 290 ZIP Codes. Orange County is the third most populated in California with over 3 million residents. The county covers 791 square miles of land and is the smallest county by area in Southern California. It is partitioned into 88 separate ZIP Codes.

Demographics of LA County and Orange County

The 2020 US Census estimates that the Los Angeles County population was 10,014,009. White (non-hispanic) account for 26.1% of Los Angeles County. The median household income (in 2019 dollars), 2015-2019 was $68,044. The percentage of persons aged 25 years or greater with a high school graduate degree or higher, 2015-2019 was 79.1%. The 2020 US Census estimates that the Orange County population was 3,186,989. White (non-hispanic) account for 39.8% of Orange County. The median household income (in 2019 dollars), 2015-2019 was $90,234. The percent of persons aged 25 years or greater with a high school graduate degree or higher, 2015-2019 was 85.5%.

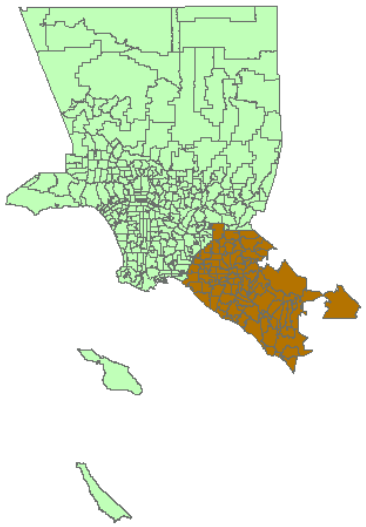


Figure 3. LA County and Orange County. Both counties are divided by ZIP Code. LA county (green) and Orange County (brown).

Data Set Construction

Rates of asthma ED visits by ZIP Code

The data source for asthma ED visits was the Emergency Department and Patient Discharge Datasets from the State of California, Office of Statewide Health Planning and Development (OSHPD). Tracking California processed OSHPD's data to calculate age-adjusted rates of asthma ED visits for California ZIP Codes by year (Tracking California, 2021). The dataset contained rates per 10,000 residents of asthma emergency department visits over the age of 18 using hospital coding. Asthma-related visits considered only primary discharge diagnosis codes. Rates were calculated from the total number of Asthma ED Visits per ZIP Code population. The primary dataset includes emergency department visits from all

licensed hospitals in California. A mean annual asthma ED visit rate for each ZIP Code was calculated by averaging the rates for the three years– 2013-2015. Rates of asthma ED visits for ZIP Codes in LAC and OC were then selected for this analysis. The Emergency Department and Patient Discharge Dataset reports patient residence only at the ZIP Code level. Individual counts or population data were not available.

The Third Unregulated Contaminant Monitoring Rule (UCMR3) Dataset

PFOA/S concentrations in drinking water for LAC and OC water districts from 2013-2015 was obtained from the UCMR3 dataset. UCMR1 emerged from the 1996 Safe Drinking Water Act (SDWA). The SDWA amendments require that every 5 years the EPA issue a list not to exceed 30 unregulated contaminants to be monitored by public water systems. The data generated from this monitoring provides a scientific basis for future regulatory actions geared towards protecting public health. The UCMR3, the third installation, was published on May 2, 2012. UCMR3 required monitoring for 30 contaminants between 2013 and 2015 among PWS serving more than 10,000 people.

The largest sources of surface water supply in LA County come from the Los Angeles River and the San Gabriel River. The coastal plain of Los Angeles groundwater basin is the primary source of groundwater. In Southern California, the suspected PFAS sources for water systems are waste-water treatment plants located upstream of managed aquifer recharge (MAR) operations. Results from UCMR3 indicated that PFAS contamination in LA County is present in 10 water districts:

1. California Water Service Company - East Los Angeles
2. City of Downey Water Division

3. Golden State Water Company - Norwalk
4. Liberty Utilities - Bellflower/ Norwalk
5. Montebello Land And Water Company
6. Orchard Dale Water District
7. City of Pico Rivera Water Division
8. Pico Water District
9. Santa Clarita Water Division
10.Valencia Water Company

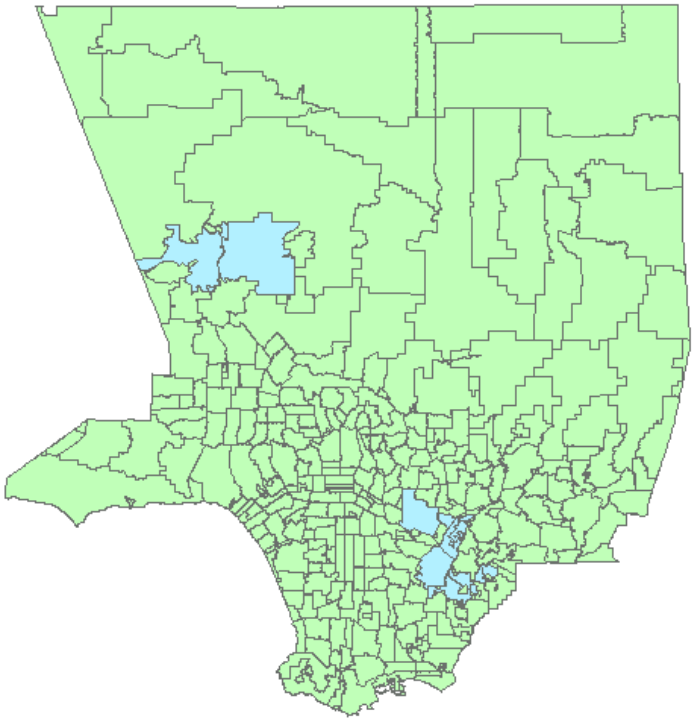


Figure 4. LA County PFAS contaminated water districts. Water districts with detected PFAS levels (blue).

The northern portion of Orange County lies above a large aquifer known as the Orange County Basin. This water source provides a significant portion of water for the Orange County cities north of Newport Beach and Irvine. It is estimated that 2.4 million OC residents rely on the Orange County Basin for half their water needs. Results from UCMR3 indicated that PFAS contamination in Orange County is present in 5 water districts:

1. City of Anaheim Water Division
2. City Of Fullerton Water Division
3. City of Garden Grove Water Division
4. City of Orange Water Division
5. Yorba Linda Water District

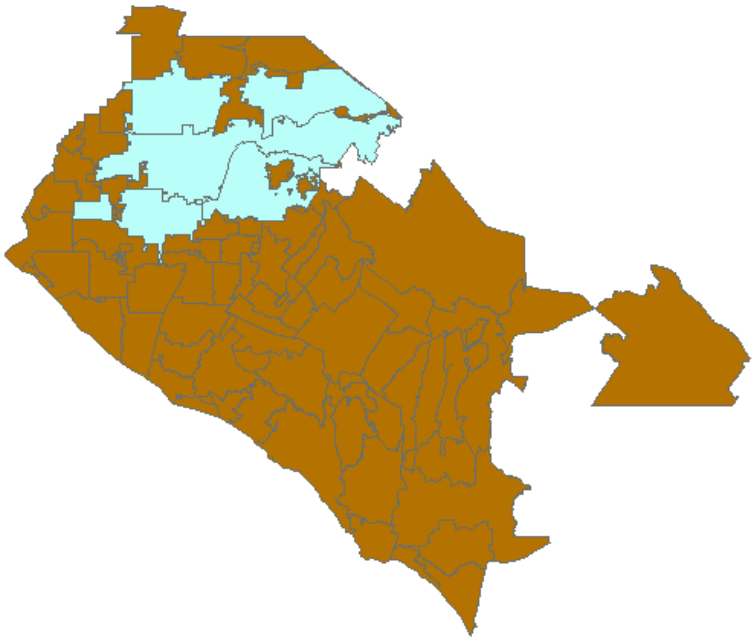


Figure 5. Orange County PFAS contaminated water districts. Water districts with detected PFAS levels (blue).

PFOA/S derived serum concentrations by water district

Available measures of drinking water PFOA/S in the UCMR3 data for LAC and OC water districts were averaged for 2013-2015. These measures were used to derive average PFOA/S serum concentrations for the contaminated LA and OC water district. Using a basic pharmacokinetic (Pk) model that assumes typical background serum levels from non-drinking water exposure sources (https://www.ics.uci.edu/~sbartell/pfascalc.html), expected serum concentrations resulting from long-term consumption of PFOA/S

contaminated drinking water were calculated at measured concentrations above the reporting limit within each water district. Only ZIP Codes with contaminated water and adjacent ZIP Codes were included in the analysis. Modeled serum concentrations were used instead of estimated water levels for easier comparison to other populations such as NHANES.

Comparison of rates of asthma ED visits to PFOA/S concentrations

To investigate a potential association between rates of asthma ED visits and PFOA/S levels a common unit of location needed to be obtained as the original datasets used ZIP Codes for asthma ED visits and water districts for PFOA/S concentrations. To resolve this conflict, PFOA/S concentrations were assigned to ZIP Codes based on the location of the water districts. ArcGIS was used to determine the land area percentage that distinct water districts covered within a given ZIP Code. A weighted average based on these percentages was used to determine the PFOA/S concentration within a given ZIP Code.

ArcGIS

Data set included ArcGIS shapefiles of geographic areas of ZIP Codes and water districts of Los Angeles County (LAC) and Orange County (OC). Each county had a separate shapefile for ZIP Code and water district. Shapefiles were imported into ArcGIS. ZIP Code shapefiles were in a projected coordinate system in units of feet. Water district shapefiles were originally in a geographic coordinate system but were immediately projected into the same coordinate system as ZIP Code shape files after import to ArcGIS. The tabulate intersect analysis was used to determine the percent of land that a respective water district covers

within a given ZIP Code. To accomplish this, ZIP Code shapefiles were used as the 'zone' and water district shapefiles were used as the 'class'. This analysis was done separately for each county and a table containing area and percentage of intersection were generated.

CalEnviroScreen 4.0

Air quality data was obtained from the California Environmental Screen 4.0 (CalEnviroScreen 4.0). The CalEnviroScreen is a screening tool that helps in identifying and mapping areas in California burdened by pollution. The CalEnviroScreen was developed by the Office of Environmental Health Hazard Assessment (OEHHA) as part of CalEPA's environmental justice program. This mapping tool evaluates the magnitude of pollution from multiple sources including environmental, health, and socioeconomic information. Scores are generated for every census tract in California, and these are ranked based on data from state and federal government sources. Data for ranking include prevalence of pollutants, poor environmental conditions, socioeconomic status, and abundance of local health conditions. Higher scores represent census tracts with higher amounts of pollution burden. Census tract data is used with 21 indicators to create a single score. The Census tract data was aggregated into ZIP Code data using the provided ZIP Code metrics. The CalEnviroScreen ZIP Code estimates were assigned to 2010 census blocks using areal apportionment. Areal apportionment is a simple technique to divide and allocate values by area. The area of each Census Tract-Zip Code intersection is calculated, and then census tract values are given to zip code boundaries proportional to how much of its area lies within each zip code. Indicators of interest for this study were $PM_{2.5}$ and diesel particulate matter. The California Air Resources Board (ARB) collected estimated emissions from

several sources, primarily from years 2012-2019. CalEnviroScreen 1.0 was released in 2010. Up-to-date date data and contributions from governmental organizations have been incorporated in subsequent versions of the tool. CalEnviroScreen 2.0 was released in 2014 and 3.0 in 2017. The CalEnviroScreen 4.0 was released in February 2021. Data was acquired from the CalEnviroScreen 4.0 data set in July 2021. The CalEnviroScreen Version 4.0 receives periodic updates as data becomes available and as methodologies improve.

Census Data

Demographic information was obtained from the United States Census Bureau (census.gov). The retrieved data was in ZIP Code Tabulation Area (ZCTA) format, equivalent to a ZIP Code. ZCTAs are based on aerial representations of United States Postal Service ZIP Code service areas. Retrieved information included race, education, and income from the 2020 United States census.

Joining the Data

Intersection tables generated in ArcGIS were exported into Excel files. LAC and OC Excel files were merged together. PFOA/PFOS concentrations by water district were added to the Excel file. PFOA/PFOS was calculated for each ZIP Code by adding up the weighted PFOA/ PFOS concentration of the water districts that made up a given ZIP Code and weighting each of the component concentrations by land area percentage that the water district covered in the given ZIP Code. Unexposed ZIP Codes were assigned the background concentrations used in the PK modeling. PFOA background concentration is 2 ng/mL.

PFOS background concentration is 5 ng/mL. Asthma ED rates were then linked by ZIP Code. There are a total of 75 ZIP Codes in this analysis.

Examining the Data

Confounders

Literature-based discovery was used to identify potential confounding variables in the study. Identified variables include: Diesel, Particulate Matter, Race, and Education. Air pollution data for $PM_{2.5}$ and diesel were obtained from CalEnviroScreen 4.0 (https://calenviroscreen-oehha.hub.arcgis.com/#Data). Income/socioeconomic status (SES), race, and education data was obtained from United States Census Information. The parameter that was assessed for income/SES was median household income. The parameter that was assessed for race was percentage of non-Hispanic White. The parameter that was assessed for educational attainment was percentage of less than high school Diploma.

Exploratory analysis

Exploratory data analysis is a statistical approach that focuses on looking at data to see what it seems to say as opposed to analytical approaches that focus on testing pre-defined hypotheses and require a well-defined question to ask from the data (Hoaglin, 2003). Using trial-and-error as part of a process of uncovering hidden features of the raw data, it is possible to discover structure, find patterns, and identify relationships. In this project, exploratory analysis was conducted by summarizing the data frame in R and visualizing the data with a number of graphs, including histogram plots of PFOA/S levels. Scatterplots

were graphed measuring Asthma ED Visits and PFOA/S concentrations. In addition, several maps were constructed to visualize the distribution of PFOS/PFOA geographically.

Correlation Analysis

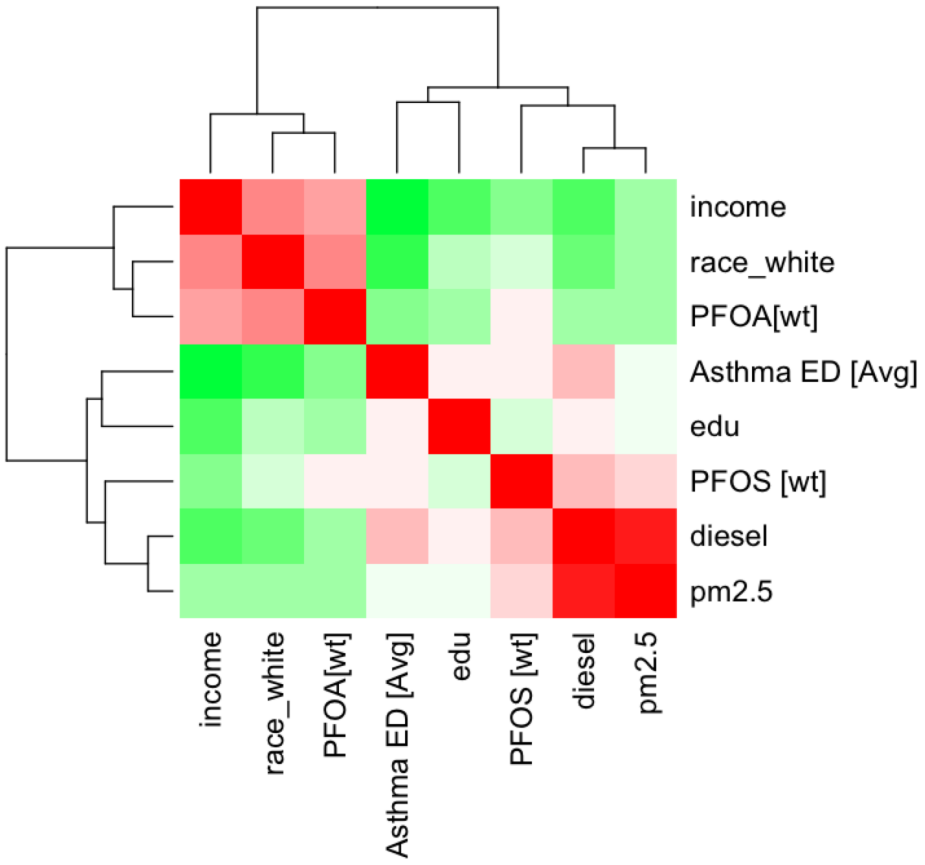


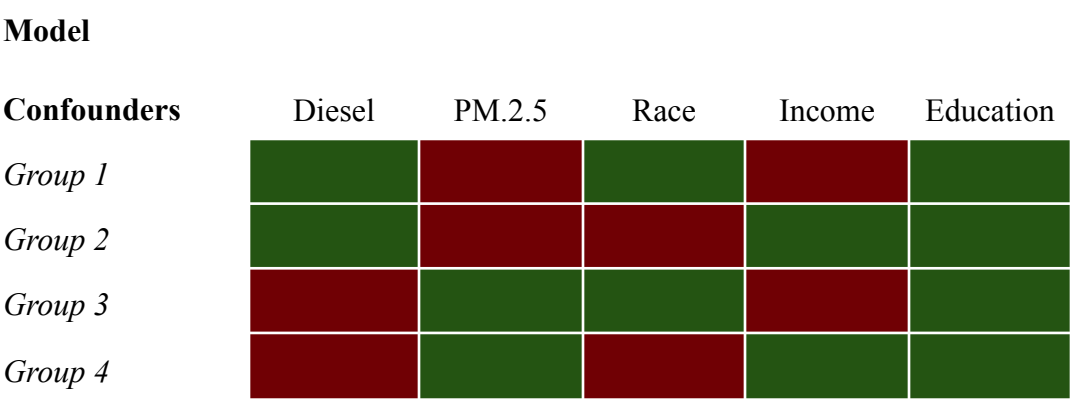
Figure 6. Heat Map of Variables. Red areas represent a stronger correlation. Green areas represent weaker correlation.

To understand the interrelationship among predictor variables, correlation analysis was conducted and a heat map of the results was constructed. Correlation coefficients are used to assess the magnitude and direction of the linear relationships between pairs of variables. High correlations are closer to 1 and correspond to red on the heat map. Low correlations are closer to 0 and correspond to Green on the heat map. Correlation analysis was not used

as a stand alone solution, but instead was accompanied by the regression analysis and primarily used as a tool to uncover any collinearity between the study's variables.

Spearman correlation analysis was used to investigate socioeconomic (race, income, education) and pollutant variables (PFOS, PFOA, Diesel, $PM_{2.5}$) along with asthma ED rates. Spearman's correlation coefficient rho (ρ) is non–parametric in nature and can be used when the data doesn't follow the normal distribution. The value ρ =1 indicates a perfectly positive correlation. On the other hand ρ = -1 suggests a perfectly negative correlation. The value ρ =0 suggests no correlation at all. Strong relationships were found between the pollutants Diesel and $PM_{2.5}$ as well as the socioeconomic factors race and income. Based on the results of the correlation analysis, combinations of the initial confounding variables (Table 1) were broken into several groups and separately analyzed.

Table 1. Collinearity-adjusted Model Combinations. Green represents confounding variables that were included in each model. Red represents excluded variables.



Regression Analysis

Multiple Linear Regression

Associations across affected LA and OC ZIP Codes were explored by developing a multiple regression model with rates of Asthma ED visits as the dependent variable and PFOA/S concentrations as the primary predictor variables. The model also controlled for air quality and SES factors. Multiple linear regression models the relationship between dependent and independent variables and is formally represented as follows:

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$$

Where y is the dependent variable, α_1, α_2 ,..., α_n are regression coefficients, and x_1, x_2..., x_n are independent variables. The main features of regression coefficients are sign, magnitude, and significance (Eberly, 2007). The positive or negative sign of the regression coefficient for an independent variable indicates the direction of the association with the dependent variable. The magnitude of each coefficient quantifies the observed effect between the independent variable and the dependent variable. Significance is the observed result based on sample data which represents the entire population.

In the multiple linear regression, the model was fit using ordinary least squares (OLS). Linear least squares estimates the unknown variable by minimizing the sum of the squares of the differences between actual values and predicted values (Eberly, 2007). The error of the model is defined using a loss function. Residuals are differences between actual and predicted values of each target variable. Residuals are calculated using the following formula:

$$E_i = y_i - f(x_i)$$

Where yi is the actual value and $f(xi)$ is the predicted value based on the multiple linear regression function.

The regression models can be evaluated using regression metrics such as R^2 score and Mean squared error. The R^2 coefficient gives a measure of how accurately future samples are likely predicted by the model (Miles, 2005). The best possible score is 1.0 and indicates that the model perfectly fits the data. Scores can be negative if the model does not follow the trend of the data. A model that always predicts the expected value of y regardless of the input variables, such as a constant y=1 function, has R^2 score zero.

Although multiple linear regression is a simple and robust technique, it does have pitfalls such as not generalizing well for nonlinear data and performing poorly in the presence of collinearity between variables. Several regression methods have been developed to overcome these limitations, including linear regression with categorical variables and spline regression, which were also performed as part of the study.

Linear Regression with Categorical Variables

When the magnitude of several regression coefficients of continuous data implied a strong effect in the model, the data were transformed into quartile values for modeling. The goal of categorizing the data is to reduce error and improve the predictive value of the

regression model. The process used to categorize the data into a smaller range was a quartile transform.

Quartiles describe a division of individual dataset values into four defined intervals based on a comparison to the entire set of observations. To derive quartiles, the vector of each variable must first be arranged in ascending order. Each quartile contains around 25% of the total vector values. The first quartile, Q1, contains the smallest 25% of values. On the other hand, Q4 contains the largest 25% of values. The resulting dataset was then re-evaluated using multiple linear regression with OLS.

Spline Regression

When the association with the covariate and ED data appeared non-linear, the data were also analyzed using a spline technique. Splines are piecewise lines or low-degree polynomials that are inset continuously between values, termed knots (Perperoglou et al., 2019). There are multiple types of splines, which reflect different choices in polynomials and knots. Spline regression is an approach for fitting and smoothing a line's twists. To compute the spline results, the "mgcv" package in R was used.

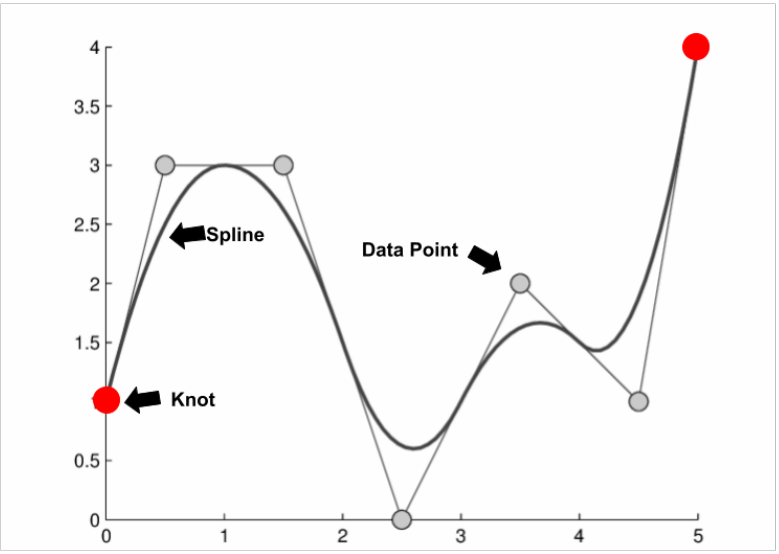


Figure 7. Sample Spline and Knot.

Chapter 4: Results

Exploratory Analysis

Descriptive Statistics

The background serum concentration for PFOA was 2 ng/mL and the background serum concentration for PFOS was 5 ng/mL. These represent exposure from the non-drinking water PFOA/S found in the environment, i.e. pots/pans, product packaging, and fire-resistant chemicals. These values were assigned to ZIP Codes in which no drinking water PFOA and/or PFOS contamination was identified. Thus, these background values represent the minimum PFOA/S in this data set (Table 2). Of the 75 ZIP Codes in the analysis, 31 had no detectable PFOA in the drinking water and were assigned serum levels of 2 ng/mL. 11 ZIP Codes had no detectable PFOS in the drinking water and were assigned serum levels of 5 ng/mL. The abundance of the undetectable PFOA/S in many water districts resulted in the background values playing a significant role in our data set to the extent where the minimum (0 percentile) and 1st quartile (25th percentile) were very close to the same value for PFOS. In the case of PFOA, 2 ng/mL was the value from the minimum (0 percentile) all the way to the 41st percentile (Table 2). The range (Min-Max) for PFOA/S in the data set was narrow with the maximum values only reaching slightly above 2x the minimum values.

Table 2: Descriptive Data for LA County and Orange County combined. Annual asthma ED visits per 10,000 by ZIP Code. Predicted serum PFOA/S (ng/mL) levels in communities with contaminated drinking water by ZIP Code.

Summary Data for PFOA/S (75 LAC and OC Zips)

	Min.	1st Qu.	Median	Mean	3rd Qu	Max
Asthma ED Visits	*28.070*	*45.670*	*72.570*	*68.860*	*83.400*	*117.330*
PFOS (ng/mL)	*5.000*	*5.011*	*5.606*	*7.381*	*10.093*	*11.299*
PFOA (ng/mL)	*2.000*	*2.000*	*2.003*	*2.550*	*3.130*	*4.448*

Descriptive data for potential confounding variables are shown in Table 3. The range of PM 2.5 was 0.0095 to18.49 μg/m^3, the median was 4.26. The range of diesel was 0.000041 to 0.61kg, the median was 0.064. There was a much larger variation in air quality between ZIP Codes with both $PM_{2.5}$ and diesel compared to the variation and PFOA/S between ZIP Codes. The range for income was $21,964 to $169,125; the median was $73,187. The percentage of white non-Hispanic in each ZIP Code ranged from 16% to 84%, the median was 63%. The percentage of individuals who did not finish high school in each ZIP Code ranged from 0% to 26%, the median was 9%.

Table 3. Descriptive data for LA County and Orange County combined by ZIP Code. PM 2.5 (μg/m^3), Diesel (kg), Income, Race (% White non-Hispanic), Education (% high school incompletion) by ZIP Code.

Summary Data for Potential Confounders: PM 2.5, Diesel, Income, Race, Education

	Min.	1st Qu.	Median	Mean	3rd Qu	Max
PM2.5	*0.010*	*2.182*	*4.256*	*4.927*	*6.886*	*18.490*
Diesel	*0.000*	*0.028*	*0.064*	*0.106*	*0.143*	*0.612*
Race	*16.00%*	*49.50%*	*63.00%*	*60.07%*	*75.00%*	*84.00%*
Income	*$21,964*	*$63,147*	*$73,187*	*$82,591*	*$101,237*	*$169,125*
Education	*0.00%*	*7.00%*	*9.00%*	*10.25%*	*12.00%*	*26.00%*

PFOA reached a maximum serum level of ~4.5 ng/mL with only 25% of the ZIP Codes having a PFOA serum level above 3 ng/L (Table 2 and Figure 8). PFOS reached a maximum serum level of 11.3 ng/mL with 25% of the ZIP Codes having at least 10 ng/ml (2x the minimum value) (Table 2 and Figure 9). On the other hand, the range of the outcome asthma ED visits that our study measures was much larger compared to the range of the risk factors PFOA/S. The minimum number of ED visits was 28.07, whereas the maximum was 4x larger at 117.33. The top 25% of ZIP Codes with regards to asthma ED visits were 3 to 4x larger than the minimum.

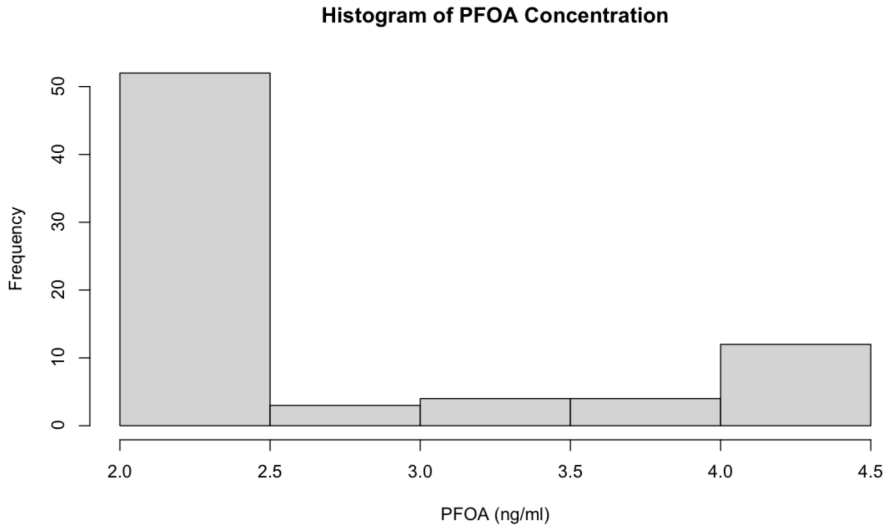


Figure 8. Histogram of Predicted serum PFOA Descriptive data for LA County and Orange County combined by ZIP Code.

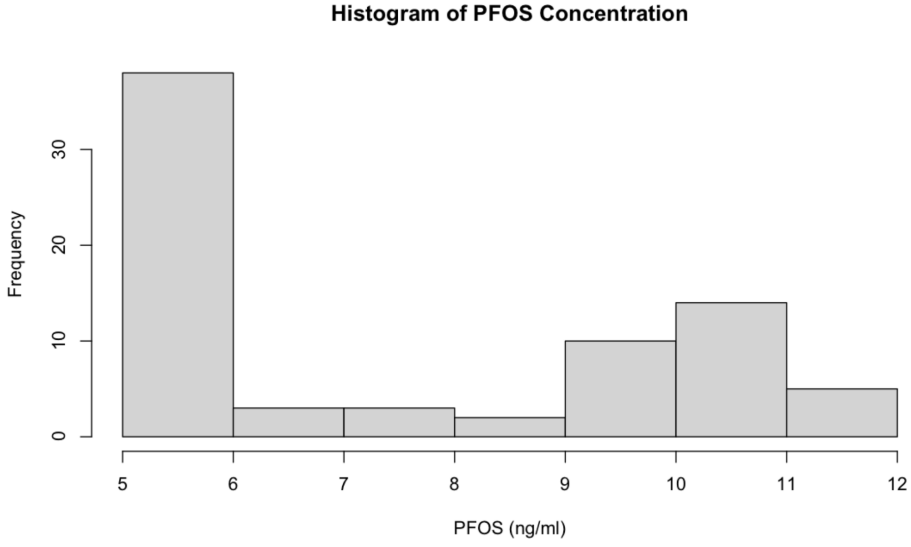


Figure 9. Histogram of predicted serum PFOS descriptive data for LA County and Orange County combined by ZIP Code.

Scatter plots were created to determine if there was a correlation between our risk factors of interest (predicted serum PFOA/S) and our outcome of interest (Asthma ED visits). Per scatter plot, there was no correlation between predicted serum PFOA and the number of asthma ED visits (Figure 10). There was also no correlation between predicted serum PFOS and the number of asthma ED visits (Figure 11). A visual analysis using ArcGIS was also used to determine if there is a correlation between predicted serum PFOA/S and Asthma ED visits. Per mapping, there was no correlation between predicted serum PFOA and the rate of asthma ED visits (Figure 12). There was also no correlation between predicted serum PFOS and the rate of asthma ED visits (Figure 12).

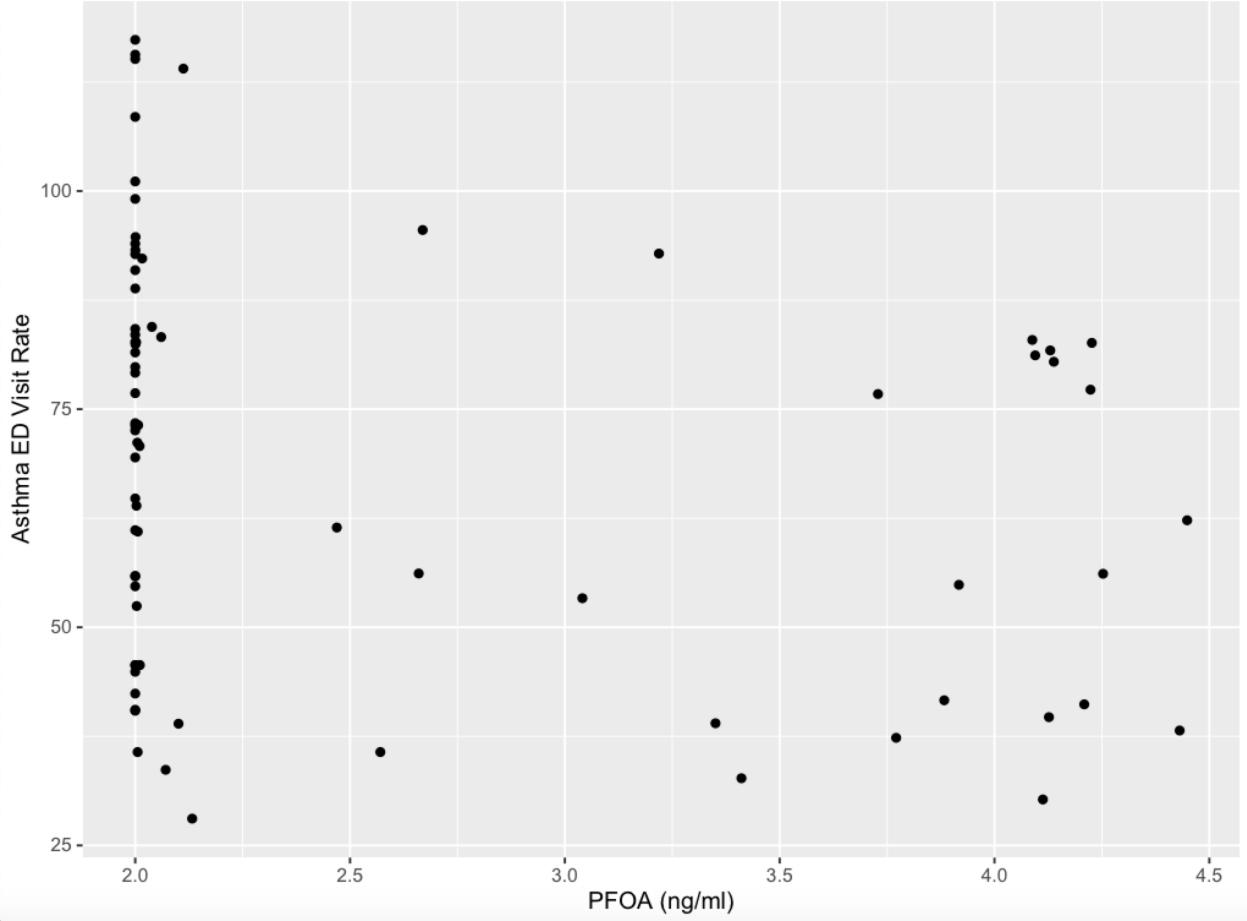


Figure 10: Scatter plot of predicted serum PFOA (x-axis) levels (ng/mL) and rate of asthma ED visits (y-axis)

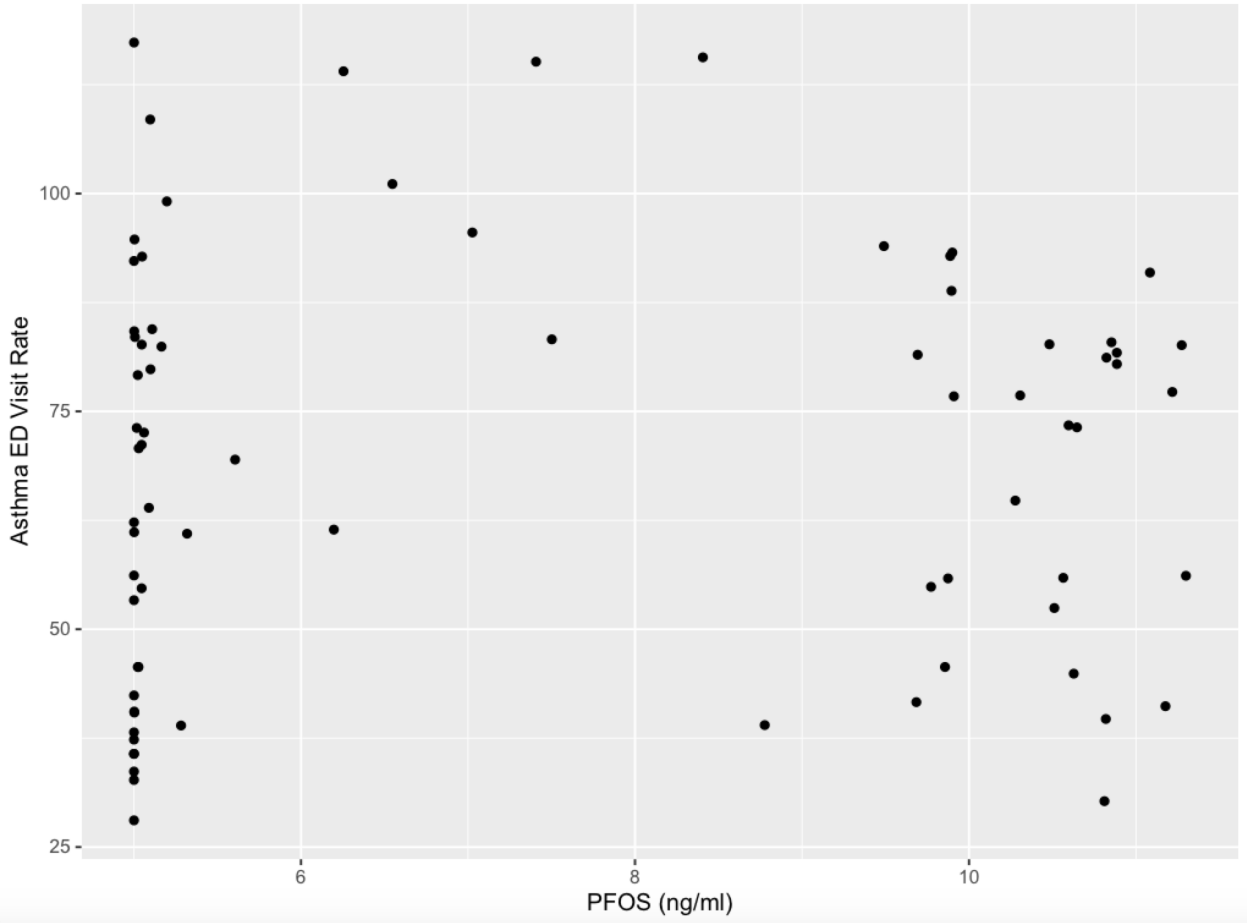
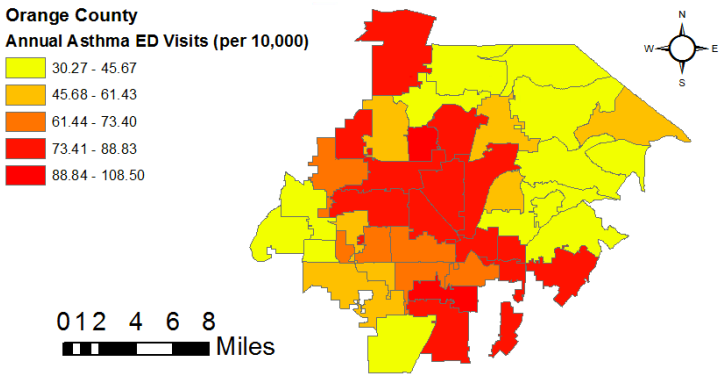


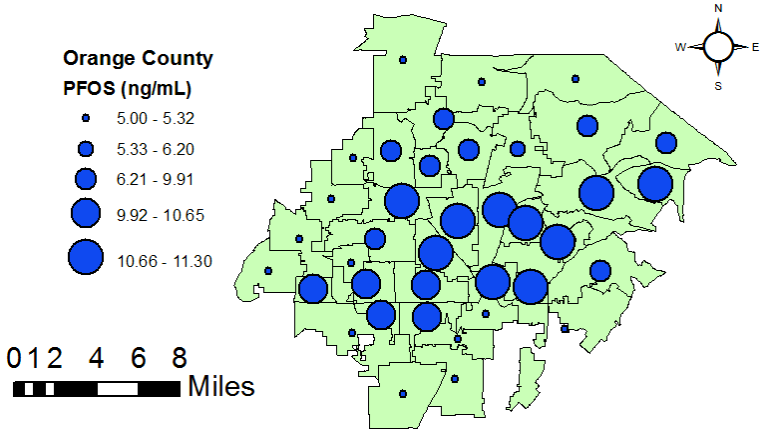
Figure 11: Scatter plot of predicted serum PFOS (x-axis) levels (ng/mL) and rate of asthma ED visits (y-axis)

Spatial Analysis

Data, including predicted serum PFOS/PFOA levels and Asthma ED Visit rates, were mapped and analyzed at the ZIP Code level for the ZIP Codes included in the study for each county. Results demonstrated distinct spatial patterns of asthma ED visit prevalence across

both LAC and OC. While there was some overlap between ED visit rate and PFAS hot spots, there was no definite pattern, suggesting that varied spatial processes are at play.





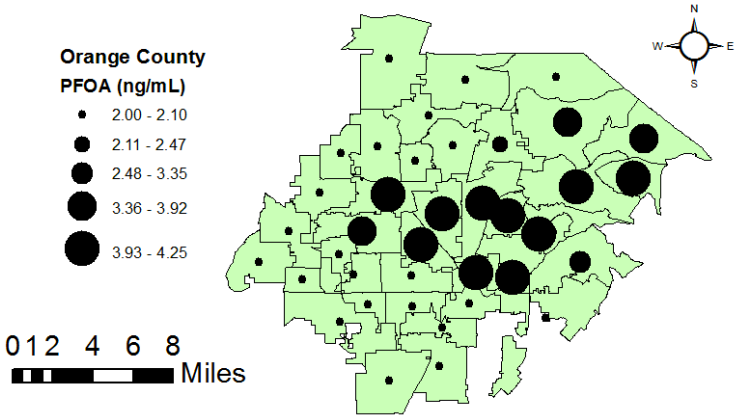


Figure 12. Orange County ArcGIS Map of Rates of Asthma ED Visits, PFOA, and PFOS in contaminated ZIP Codes. Top map: Rates of Asthma ED visits represented by graduated colors - yellow, orange, red (lowest to highest). Middle map: Predicted serum PFOS (blue) levels represented by graduated shapes - blue circles smallest to largest (lowest to highest concentration of PFOS). Botton map: Predicted serum PFOA (black) levels represented by graduated shapes - black circles smallest to largest (lowest to highest concentration of PFOA).

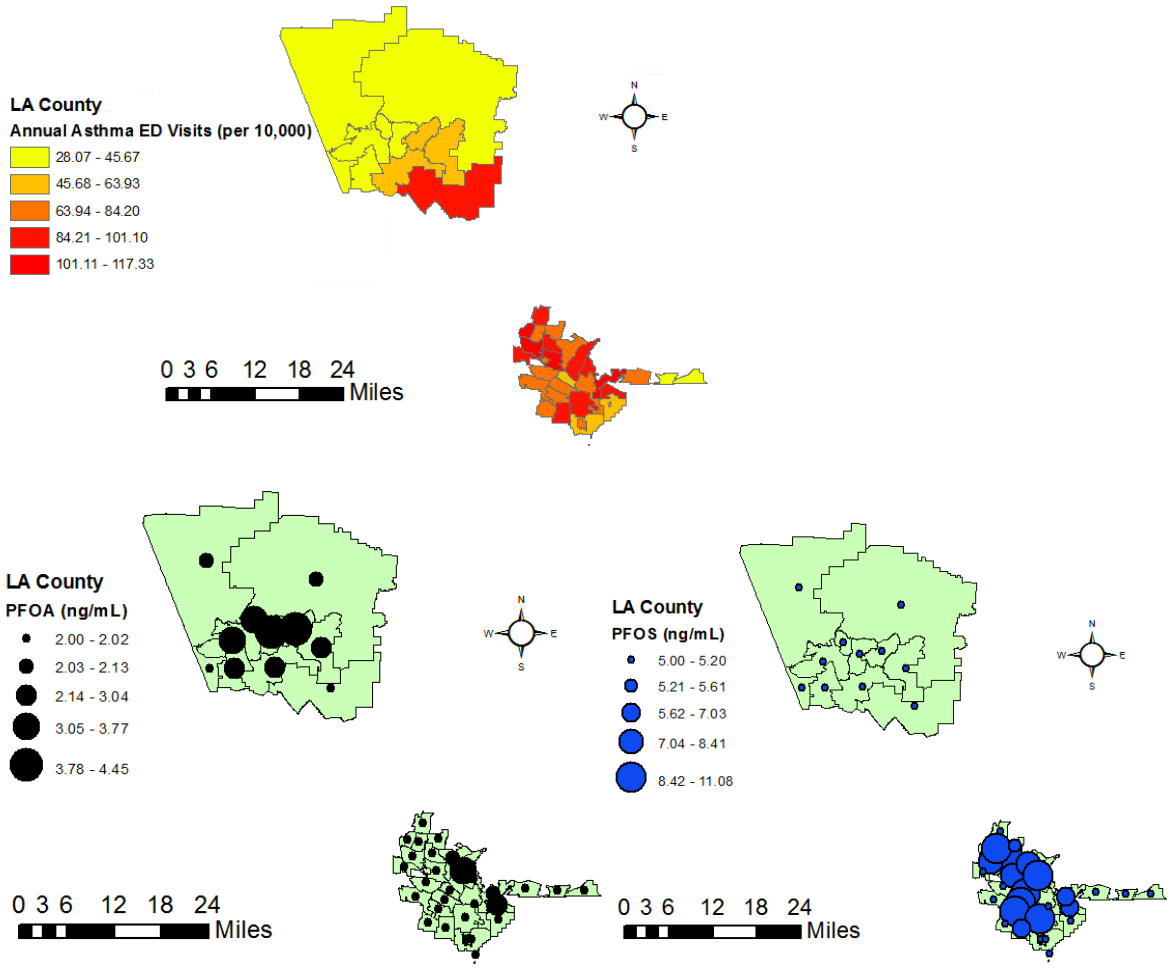


Figure 13. Los Angeles County ArcGIS Map of Rates of Asthma ED Visits, PFOA, and PFOS in contaminated ZIP Codes. Top map: Rates of Asthma ED visits represented by graduated colors - yellow, orange, red (lowest to highest). Middle map: Predicted serum PFOS (blue) levels represented by graduated shapes - blue circles smallest to largest (lowest to highest

concentration of PFOS). Botton map: Predicted serum PFOA levels (black) represented by graduated shapes - black circles smallest to largest (lowest to highest concentration of PFOA).

Regression Analysis

Simple Linear Regression

As part of exploratory analysis, a simple regression was done with predicted serum PFOS as the only independent variable. The relationship was not statistically significant and the model had quite a low R-Squared of 0.004. Likewise, PFOA was evaluated as a single independent variable, and the model had an R-squared value of 0.06. Therefore, further analysis that included additional explanatory variables was an appropriate next step.

Table 4. Result of a single linear regression of PFOS and single linear regression of PFOA versus rate of Asthma ED visits

Variable	Estimate	Confidence Interval
PFOS (Only)	0.61	-1.47 , 2.69
PFOA (Only)	-6.78	-12.74, -0.82

Multiple Linear Regression

Linear regression analysis was performed to determine if there is a correlation between our risk factors of interest (predicted serum PFOA/S) and our outcome of interest (rate of Asthma ED visits). Continuous serum PFOS concentrations showed a positive association with rates of asthma visits whereas serum PFOA showed a negative association with rates

of asthma ED visits. Potential confounding variables are also included in the regression analysis. Diesel and education were associated with increased Asthma ED visit rates. $PM_{2.5}$, income, and race were associated with decreased asthma ED visit rates (Table 6).

Table 5. Result of a single linear regression of independent continuous variables of interest and potential confounding variables used to predict Asthma ED visits.

Variable	Estimate	Confidence Interval
PFOS	1.67	-0.23 , 3.57
PFOA	-6.53	-12.21 , -0.85
PM 2.5	-10.22	-18.06 , -2.38
Diesel	12.68	5.49 , 19.87
Income	-14.05	-22.71 , -5.39
Race	-2.79	-11.06 , 5.48
Education	8.14	0.71 , 15.57

Simple Linear Regression with Categorical Variables

To further examine negative associations (i.e., PFOA and $PM_{2.5}$) with Asthma ED visits, the variables were transformed into quartiles and categorical analysis was performed (Table 7). As the concentration of serum PFOS increased from Q2-Q3, there is an increase in the positive regression coefficients with asthma ED visit rates. Surprisingly, in the 4th quartile of PFOS there was a decrease in the rate of Asthma ED visits suggesting that the relationships may not be linear. Unexpectedly as the concentration of PFOA increased from Q2-Q4, there is a decrease in the positive association and an increase in the negative

association with to asthma ED visit rates. Unexpectedly, as the concentration of $PM_{2.5}$ increased from Q2-Q4, there was an increase in the negative association with asthma ED visit rates. On the other hand, diesel produced an estimate in a tight range (9-10) for each quartile. With increasing income from Q2-Q4, a decrease in Asthma ED visit rates was observed. Race was defined as non-Hispanic White. As expected, as the number of non-Hispanic White increases from Q2-Q4, there is a decrease in the positive association and an increase in the negative association with asthma ED visit rates. Education is defined as percent incomplete high school. Likewise as expected, as the percentage of individuals who did not complete high school increased from Q2-Q4 in a ZIP Code there was an increased asthma ED visits.

Table 6. Results of quartile linear regression model of independent variables of interest and potential confounding variables used to predict rates of asthma ED visits.

Variable	Estimate	Confidence Interval
PFOS_Q2	0.32	-11.77,12.40
PFOS_Q3	10.11	-3.87,24.09
PFOS_Q4	-2.19	-16.37,11.98
PFOA_Q2	0.73	-11.55,13.01
PFOA_Q3	0.11	-12.62,12.84
PFOA_Q4	-1.79	-16.37,12.79
pm2.5_Q2	-3.39	-22.52,15.75
pm2.5_Q3	-4.04	-26.11,18.03
pm2.5_Q4	-12.09	-34.97,10.78

diesel_Q2	9.14	-11.37,29.66
diesel_Q3	10.29	-13.10,33.70
diesel_Q4	10.15	-15.58,35.88
income_Q2	-4.47	-16.36,7.42
income_Q3	-16.28	-31.77,-0.79
income_Q4	-31.69	-47.33,-16.05
race_Q2	4.15	-9.27,17.56
race_Q3	1.12	-12.37,14.61
race_Q4	-11.20	-27.59,5.19
edu_Q2	-0.72	-13.27,11.84
edu_Q3	-7.53	-18.60,3.54
edu_Q4	5.29	-6.81,17.38

Correlation Analysis

After identifying which independent variables were associated with the outcome , correlation analysis was conducted to understand the interrelationship among predictor variables. Spearman correlation analysis was used to investigate socioeconomic (race, income, education) and pollutant variables (serum PFOS, serum PFOA, Diesel, $PM_{2.5}$) along with asthma ED rates. A table representing numerical values of the correlation analysis shows positive and negative values ranging from 0 to 1 (Table 4). For example, $PM_{2.5}$ and diesel show a value of 0.84 on the table. This suggests a strong positive relationship between the two variables. Conversely, race and $PM_{2.5}$ have a low correlation value of -0.18 on the table that is close to 0. Positive signs represent a positive correlation whereas negative signs represent an inverse correlation. Strong relationships were found between

the pollutants Diesel and $PM_{2.5}$ as well as the socioeconomic factors race and income (Table 4). We ran four different models to compare combinations: 1) race and diesel, 2) race and $PM_{2.5}$, 3) income and diesel, 4) income and $PM_{2.5}$. The model that included income and $PM_{2.5}$ had the greatest predictive value.

Table 7. Correlation Coefficients. Numerical representation of heat map. Numbers closer to +1 or -1 represent stronger positive or negative correlations respectively. Numbers closer to 0 represent weaker correlations.

	PFOS	**PFOA**	**Asthma ED**	**PM2.5**	**diesel**	**income**	**race white**	**edu**
PFOS	1.00	0.16	0.22	0.29	0.34	-0.24	-0.02	-0.01
PFOA	0.16	1.00	-0.29	-0.13	-0.20	0.43	0.55	-0.19
Asthma ED	0.22	-0.29	1.00	0.10	0.32	-0.73	-0.50	0.20
PM2.5	0.29	-0.13	0.10	1.00	0.84	-0.19	-0.18	0.07
diesel	0.34	-0.20	0.32	0.84	1.00	-0.45	-0.32	0.14
income	-0.24	0.43	-0.73	-0.19	-0.45	1.00	0.54	-0.44
race_white	-0.02	0.55	-0.50	-0.18	-0.32	0.54	1.00	-0.11
edu	-0.01	-0.19	0.20	0.07	0.14	-0.44	-0.11	1.00

Categorical Regression without Collinearity

Table 8. Model Variables and R-Squared for Categorical Regression w/o Collinearity.

	PFOA	PFOS	Diesel	PM.25	Race	Income	EDU	R-Squared
Model 1	✓	✓	✓		✓		✓	0.484
Model 2	✓	✓	✓			✓	✓	0.602
Model 3	✓	✓		✓	✓		✓	0.511
Model 4	✓	✓		✓		✓	✓	0.605

Several combinations of confounders were considered as separate models and compared based on R-squared values (Table 8). While interactions were considered for combinations of race, income, diesel and $PM_{2.5}$, we only present specific results (Table 9) for the model that leveraged income and $PM_{2.5}$. This model had the strongest R-squared value at 0.605. However, there was no statistical evidence for a relationship between predicted serum PFOA/PFOS and asthma ED visit rates.

Table 9. Results of a multiple variable regression model of predicted serum PFOS/A modeled categorically and adjusted for $PM_{2.5}$, Income, and Education.

Variable	Estimate	Confidence Interval
PFOS_Q2	2.01	-9.66,13.69
PFOS_Q3	13.48	0.88,26.08
PFOS_Q4	-0.09	-13.46,13.26
PFOA_Q2	2.05	-9.76,13.86
PFOA_Q3	-0.04	-12.38,12.28
PFOA_Q4	-4.22	-17.96,9.51

pm2.5_Q2	3.83	-7.73,15.40
pm2.5_Q3	2.53	-9.50,14.57
pm2.5_Q4	-3.40	-15.31,8.51
income_Q2	-2.32	-13.84,9.18
income_Q3	-17.22	-31.07,-3.38
income_Q4	-35.70	-49.48,-21.91
edu_Q2	-2.23	-14.24,9.77
edu_Q3	-5.88	-16.51,4.75
edu_Q4	5.92	-5.78,17.63

Spline Regression

As previously stated, some results from linear regression analysis of quartiled data seem to indicate that the relationship between dependent variables and outcome may be non-linear. For this reason, we ran a spline regression analysis (Table 10). PFOS, PFOA, diesel, income, and education were used in the spline model. These variables were all modeled using spline terms. In this initial spline analysis, only income was significant and the R-sq.(adj) = 0.533.

Table 10. Spline Regression Analysis

Variable	Estimate	Confidence Interval
s(PFOS_Q)	1.71	-4.17,7.59
s(PFOA_Q)	2.04e-11	-7.84,7.84
s(diesel_Q)	1.06e-10	-7.84,7.84

s(income_Q)	2.81	-3.07,8.69
s(edu_Q)	0.90	-4.98,6.78

The spline analysis was repeated by modeling PFOA and PFOS with splines and the other covariates modeled categorically by quartiles. This second spline analysis resulted in a higher adjusted R-squared of 0.541 (Table 11). Again the only significant variable is income.

Table 11. Mixed Spline Regression Analysis

Variable	Estimate	Confidence Interval
s(PFOS_W)	3.19	-4.494,10.876
s(PFOA_W)	1.06	-1.133,3.257
pm2.5_Q2	4.61	-6.52,15.74
pm2.5_Q3	1.01	-10.33,12.34
pm2.5_Q4	-0.01	-11.34,11.33
income_Q2	0.85	-10.46,12.16
income_Q3	-17.32	-29.6,-4.99
income_Q4	-34.52	-46.79,-22.26
edu_Q2	-1.66	-12.61,9.29
edu_Q3	-6.56	-16.55,3.43
edu_Q4	3.91	-7.28,15.10

Spline data was plotted for serum PFOS vs rates of asthma ED visits. In the figure, PFOS values range from 5 to 11. At the lowest concentration of PFOS, there is a steady increase in rates of asthma ED visits which peaks at 7, and then steadily declines until 10.5.

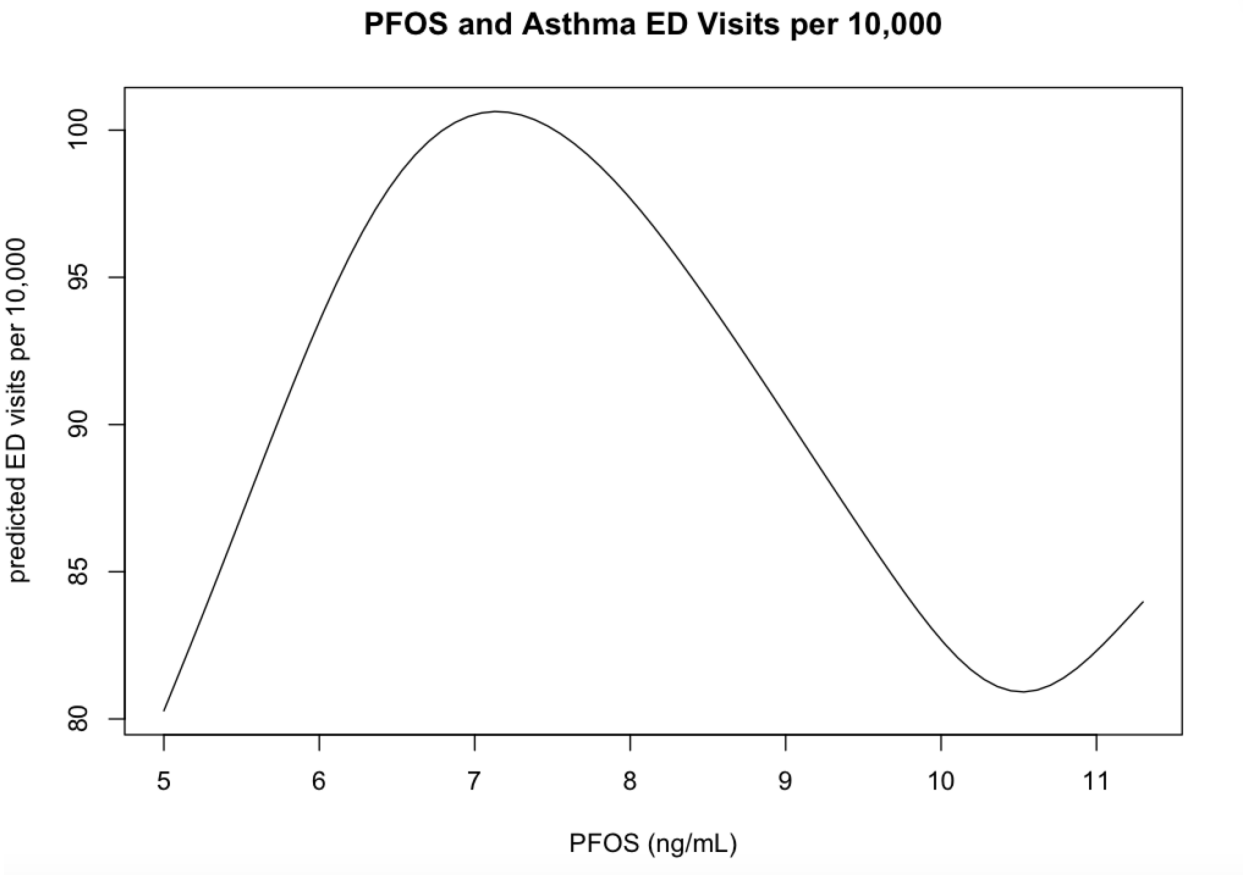


Figure 14. Predicted serum PFOS and predicted number of asthma ED visits.

Evaluation of Regression Techniques

Five different types of regression models were evaluated with several variations of each type (Table 12). Correlation Analysis revealed collinearity between several confounders and this factor was considered for final model selection along with R-squared value. Income was the variable that added the most predictability to each model. A categorical regression model with $PM_{2.5}$, education, and income as confounders was shown to be the most appropriate.

Table 12. Variable Type and R-Squared for Various Models.

Model	**Raw Variables**	**Categorical Variables**	**Smoothed Variables**	**R-squared**
Simple Linear Regression	PFOS			0.004
Multiple Linear Regression	All			0.525
Categorical Regression		All		0.649
Categorical Regression w/o Collinearity		**PFOS, PFOA, $PM_{2.5}$, Income, Education**		**0.605**
Spline Regression			All	0.533
Mixed Spline		$PM_{2.5}$, Income, Education	PFOS,PFOA	0.541

Chapter 5: Discussion

PFOA Exposure Patterns

In our study, derived serum concentrations by water districts for PFOA ranged from 2 to 4.5 ng/mL with a median of 2 ng/mL. In some cases, this level is higher than what is reported in the literature; however, the differences in background level could be due to temporal trends and/or geographic location. The mean PFOA serum concentrations for the U.S. population from the National Health and Nutrition Examination Survey trended downward from 2.08 ng/mL to 1.42 ng/mL from 2011 to 2018 (CDC, 2021). Dong et al. measured a median PFOA serum of 1.2 ng/mL in asthmatics and 0.5 ng/mL in non-asthmatics (Dong et al., 2013). Likewise Qin et al., measured a median PFOA serum of approximately 1.02 ng/mL in asthmatics and 0.5 ng/mL in those without asthma (Qin et al., 2017). In other cases, there were reports of PFOA levels that exceed our calculated levels in this study. Such as is the case for Humblet et al., who measured a median PFOA serum of approximately 4.2 ng/mL (Humblet et al., 2014).

In our study, there was no difference between LA and Orange county in terms of modeled serum PFOA concentrations by water districts. Both counties had a median for PFOA of 2 ng/mL, which reflects the study design of using 2 ng/mL as a background serum concentration. Within LA county, the ZIP Codes that had the highest PFOA concentration were two adjacent locations: 91350 (4.4 ng/mL) and 91351 (4.5 ng/mL). Within Orange county, there were ten ZIP Codes that had between 4.1 and 4.3 ng/mL (the highest in the county) of PFOA in the drinking water. These include two clusters of ZIP Codes: cluster 1

(92801, 92802, 92805, 92806, 92807, and 92808) and cluster 2 (92865, 92866, 92867, and 92868).

PFOS Exposure Patterns

In our study, derived serum concentrations by water districts for PFOS ranged from 5 to 11.3 ng/mL with a median of 5.6 ng/mL. The level of PFOS calculated in our study was lower than that published in the literature; however, prevalence of PFAS have been decreasing over time. For instance from 1999-2008, Humblet et al. measured a median PFOS of about 16.9 ng/mL (Humblet et al., 2014). Even higher concentrations were reported by Dong et al., who measured a median PFOS of about 33.9 ng/mL in asthmatics and 28.9 ng/mL in controls (Dong et al., 2013). Likewise, Qin et al. measured a median PFOS of about 30.0 ng/mL (Qin et al., 2017). On the other hand, the mean PFOS serum concentrations for the U.S. population from the National Health and Nutrition Examination Survey trended downward from 6.31 ng/mL to 4.25 ng/mL from 2011 to 2018 (CDC, 2021). Due to exposure patterns, most communities with elevated PFAS, i.e. military or industrial environments, have primarily elevated PFOS; however, given the suspected source of Southern California water system contamination, it is likely that PFOS/PFOA exposure is correlated in our study.

In our study, there was a marked difference between LA and Orange county in terms of serum PFOS concentrations by water districts. As a result of our ZIP Code inclusion method, the median for PFOS in Orange County was nearly 2 fold greater (9.9 ng/mL) than that of LA county (5.1 ng/mL). This difference resulted from 11 ZIP Codes in LA that had

no detectable PFOS in their water districts and as such the background PFOS level of 5.0 ng/mL was assigned as the PFOS level for these ZIP Codes. Whereas all ZIP Codes in Orange County had measurable levels of PFOS in their respective water districts. Within LA county, the ZIP Code that had the highest PFOS concentration was 90022 (11.1 ng/mL) followed by a ZIP Code cluster ranging from 10.3 to 10.5 ng/mL of PFOS (90240, 90241, and 90242). Within Orange county, the ZIP Codes that had the highest PFOS concentration were four adjacent locations: 92865, 92866, 92867, and 92868.

Asthma ED Visits

In our study, rates of age-adjusted Asthma ED visits, 2013-2015, for the ZIP Codes included for LA and Orange county ranged from 28 to 117 per 10,000. The mean rate of Asthma ED visits was 69 per 10,000. The mean rate of Asthma ED visits for the 73 ZIP Codes included in this study was slightly above the CDC reported national rate of approximately 60 per 10,000 ED visits with asthma as the primary diagnosis during 2013-2015 (CDC). This elevated rate of asthma ED visits in LA and Orange county may be related to the high levels of smog and other air pollutants that encompass the LA area. This is further supported by the markedly elevated median number of asthma ED visits in the LA ZIP Codes 81 per 10,000 (mean = 75 per 10,000) compared to the median of Orange County at only 61 per 10,000 (mean = 63 per 10,000). Within LA county, the ZIP Codes that had the highest rate of asthma ED visits were 90023 (115 per 10,000), 90033 (117 per 10,000), 90040 (115 per 10,000), and 90605 (114 per 10,000). Within Orange county, the ZIP Codes that had the highest rate of asthma ED visits were 92703 (108 per 10,000) followed by 92832 (93 per 10,000).

Predicted Serum PFOA/S and Rates of Asthma ED Visits

In this study, we found that there was no statistically significant relationship between predicted serum concentrations of PFOA/S and Asthma ED visit rate in communities with PFAS drinking water contamination. However, we did observe a non-statistically significant positive relationship between these two variables in which the second and third quartiles of both PFOA and PFOS had a positive association with asthma ED visits. Nonetheless, there were wide confidence intervals for these associations, so the effect is inconclusive. This positive relationship was not seen in the fourth quartile of either. Even when modeled categorically by quartiles, the overall relationship for serum PFOS and Asthma ED visits remained positive, while that for serum PFOA was negative. We may not have been able to see a statistically significant relationship in our study due to the high level of air pollution in LA county and surrounding areas which may add confounders that were difficult to control. Another factor is the strong negative relationship between asthma ED visits and household income because of the layout of LA it is possible that wealthy households may be in proximity to the higher levels of pollutants in areas more central to the city but not consume the drinking water as a result of being able to afford filtered bottled water. It's also possible that inhaled PFOS/PFOA may be more applicable route for asthma exacerbation than ingested or dermal routes of exposure. We were unable to compare our results to other studies as we could not identify any other studies in the scientific literature that evaluated this relationship.

Limitations of the Study

There are a number of limitations to the study. One of the main limitations of our study is that our analysis uses asthma ED visits as a proxy for asthma exacerbations. Many different factors -- medication compliance, lack of continuity with a primary care provider, stress -- can result in ED visits with the primary diagnosis of asthma. Many of the aforementioned variables are more common in communities with lower socioeconomic status. Thus we attempted to control for these variables by adding income and educational attainment to our model.

Certainly, results would be more valid if data about medication compliance and asthma primary care visits were included. Studies have found that compliant asthmatic patients are significantly less likely to experience exacerbation than their less-compliant asthmatic counterparts (Stern et al., 2006). Patients and/or their caregivers often have misconceptions about the ED department, including that the ED have access to more effective treatments or that ED staff are better qualified than primary care counterparts (Al-Muhsen et al., 2015). These perceptions can lead to excess low acuity visits and deflate the explanatory value of ED visits.

Another limitation is the study assumes similar use of drinking water for all individuals within a ZIP Code as well as similar use of drinking water between ZIP Codes. For instance some families may drink exclusively tap water within a given ZIP Code while other families within the same ZIP Code may drink soft drinks exclusively. Likewise, lower income ZIP Codes may drink significantly more bottled water compared to tap water. In fact, low SES groups have more negative perceptions about the quality and safety of tap water than their

more affluent or highly educated peers (Rosinger et al., 2018). Moreover, within studied water districts, customers that consume primarily bottled water average serum PFOA concentrations that are more than 50% less than their primarily tap water consuming counterparts (Bartell et al., 2010).

Additionally, while our study addressed many important confounders for asthma ED visits, some confounders were not considered in our analysis. Important known confounders include the smoking status of patients and/or their caregivers. The habit of smoking in asthmatics can lead to destructive effects in patient outcomes and effectiveness of treatment (Stapleton et al., 2011). The possibility of other unmeasured confounders causing bias cannot be eliminated.

There are general limitations to the application of ZIP Code as the geographic unit for analysis versus more precise geographies such as individual addresses and/or census tracts. While generally more accessible, ZIP code measures can have more variation within each individual unit (Thomas et al., 2006). Moreover, our study uses population data and assumes one PFAS value for each individual within a given ZIP Code rather than directly measuring serum PFAS values. In other words, all individuals from a given ZIP Code were assumed to have roughly the same serum PFOA/S values. However, past serum level studies have found variance among different demographics. For example, one community study found PFAS serum levels were higher among older compared with younger as well as male versus female residents (Graber et al., 2019). Similarly, variables such as education, race, $PM_{2.5}$, and diesel were analyzed at a ZIP Code versus an individual level.

Finally, there were limits to the statistical approach. Our regression models studied the association between the Asthma ED response variable and PFOS/PFOA exposure variables. Besides the exposure of interest, other covariates were fitted to the model in order to control for their effects on outcome. Unfortunately, there could have been misspecification of the main exposure variables and the other covariates, and this could have adversely affected tests of the association between the exposure and response (Myers 1990). Linear regression may not have been appropriate considering the extremely non-normal distribution of the PFOA and PFOA levels by Zip Code. Moreover, there was multicollinearity for some of the key variables. This problem was addressed by sorting correlated covariates into separate models; however, the problem could have further been solved by more informative data, possibly in the form of a larger sample. A larger number of observations could have also increased statistical power. A study with inadequate statistical power has a reduced chance of detecting a true effect. Moreover, low power also reduces the likelihood that a statistically significant result reflects a true effect. In our study, the confidence intervals for most of the reported variables are very wide and could be consistent with random variations.

Strengths of the Study

Despite the limitations, there are several strengths in this study. A strength of this study is that it addresses important confounders that have been shown in the literature to impact asthma ED visits. These confounders include income, race, education, and air quality ($PM_{2.5}$ and diesel). By controlling for these confounders in our regression analysis we were able to enhance the predictive power of our overall model, though it didn't further the detection of

PFOS/PFOA effects. Another strength of this study is that we performed analysis to address a possible linear relationship between dependent and independent variables as well as an analysis to address a non linear relationship between these variables. By addressing a potential non-linear relationship we were able to enhance the predictive power of our model.

Future Studies

Future studies should be designed to address the limitations of the study. One possibility is the distribution of questionnaires to individuals in the ED with the primary diagnosis of asthma. These questionnaires can address asthma medication compliance, frequency of primary care visits, BMI, use of local drinking water. Furthermore, serum PFOA/S levels can be measured for individuals in the ED with the primary diagnosis of asthma. In fact, Asthma ED visit rates could be examined in a cohort population for which biological specimens are available. Moreover, the Agency for Toxic Substances and Disease Registry (ATSDR) and the EPA have compiled a list of communities across the country that have been shown to have PFAS water contamination. It could be informative to conduct a large ecological study to compare asthma ED rates in the communities with known and likely higher PFAS contamination than a sample of control communities.

Policy implications

Significant evidence from literature review supports the belief that PFOS and PFOA are toxins that have the potential to harm human health; therefore, the implications of exposure can't be ignored. The regulation of PFOA/S has been a contentious yet salient

issue in Environmental Health. In 2006, due to public health concerns the EPA convinced eight major companies in the PFASs industry to voluntarily phase out PFOA/S by 2015 (Cousins et al., 2020). Despite this push in the right direction, there are currently no federal PFAS drinking water standards. This is concerning considering increasing toxicological and epidemiological evidence of the adverse health effects PFAS. This combined with the widespread environmental presence of PFAS makes this an even more alarming issue. PFAS are known drinking water contaminants, and there is documented population-level exposure identified in NHANES.

In 2016, the EPA developed health advisories (HAs), which are only recommendations (Post, 2021). This indicates that they are non-enforceable. Furthermore, some research has suggested that the EPA's HA are not sufficiently protective. Studies suggest that drinking water with 70 ng/L PFOA/S increases blood levels above current U.S. background levels (Cousins et al., 2020). As a result, seven states have developed guidelines for drinking water PFOA/S levels. California has adopted non-regulatory notification levels for PFOA and PFOS in drinking water. These levels are 14 ng/L for PFOA and 13 ng/L for PFOS. Water systems are not required to conduct monitoring. However, if test results exceed a specified level, water systems must comply with state public notification requirements (Pontius, 2019).

In 2018, the Agency for Toxic Substances and Disease Registry (ATSDR) determined minimal risk levels (MRLs) for intermediate duration exposure for PFOA/S. Though MRLs used similar measuring units to the EPA, the MRL values for PFOA (3 ng/kg/day) and PFOS

(2 ng/kg/day) were up to ten times lower than the metrics the EPA used to develop its health advisories (Buser, 2018) . In addition to the health advisories, the EPA's PFOA Stewardship Program includes PFOA and its precursors (Cordner et al., 2019). However, EPA has not included PFOA/S in the fourth cycle of UCMR testing. EPA-validated drinking water testing protocols exist for some PFAS (EPA Method 537), though validated methods are lacking for others and for media such as groundwater (Cordner et al., 2019).

The mixed results of this study and others suggests ongoing research of PFOA/S is still a high priority in order to reduce risk and inform future policy. Increased knowledge can help policymakers and other professionals in the field to properly address this health priority.

Chapter 6: Conclusion

PFOA/S are ubiquitous and highly stable toxins in our environment. Mounting evidence from animal and epidemiological studies has begun to elucidate the toxicological profile and health effects of these compounds. Asthma is one of the health effects that has been linked to PFOA/S.

In our study, we explored the association between predicted serum PFOA/S concentrations in communities with drinking water contamination and asthma ED visits in Orange and LA counties. These studies were conducted at the population level using data sets containing drinking water concentrations of PFOA/S and number of asthma ED visits within the ZIP Codes of the 2 counties. We conducted geographical analysis by quantitatively mapping predicted serum PFOA/S calculated using measured drinking water levels and asthma ED visits by ZIP Code. We performed statistical analysis using various regression methods to determine if predicted serum PFOA/S levels were associated with asthma ED. These statistical analyses controlled for various potential confounders of this relationship — race, education, household income, air quality (diesel, $PM_{2.5}$). Mapping and statistical analysis show no clear relationship between PFOA/S serum levels and rates of asthma ED visits.

Resources

Al-Muhsen, S., Horanieh, N., Dulgom, S., Al Aseri, Z., Vazquez-Tello, A., Halwani, R., & Al-Jahdali, H. (2015). Poor asthma education and medication compliance are associated with increased emergency department visits by asthmatic children. *Annals of thoracic medicine*, *10*(2), 123.

Arrieta, M. C., Stiemsma, L. T., Dimitriu, P. A., Thorson, L., Russell, S., Yurist-Doutsch, S., ... & Brett Finlay, B. (2015). Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Science translational medicine*, *7*(307), 307ra152-307ra152.

Azizpour, Y., Delpisheh, A., Montazeri, Z., Sayehmiri, K., & Darabi, B. (2018). Effect of childhood BMI on asthma: a systematic review and meta-analysis of case-control studies. *BMC pediatrics*, *18*(1), 1-13.

Bartell, S. M. (2017). Online serum PFOA calculator for adults. *Environmental health perspectives*, *125*(10), 104502.

Bartell, S. M., Calafat, A. M., Lyu, C., Kato, K., Ryan, P. B., & Steenland, K. (2010). Rate of decline in serum PFOA concentrations after granular activated carbon filtration at two public water systems in Ohio and West Virginia. *Environmental health perspectives*, *118*(2), 222-228.

Buekers, J., Colles, A., Cornelis, C., Morrens, B., Govarts, E., & Schoeters, G. (2018). Socio-Economic Status and Health: Evaluation of Human Biomonitored Chemical Exposure to Per- and Polyfluorinated Substances across Status. International journal of environmental research and public health, 15(12), 2818.

Buser, M. C. (2018, September). Derivation of the Agency for Toxic Substances and Disease Registry's (ATSDR's) Provisional Minimal Risk Levels (MRLs) for PFOA, PFOS, PFHXs, and PFNA. In *ISEE Conference Abstracts* (Vol. 2018, No. 1).

CDC. (2021, December 16). *Early release: Per- and polyfluorinated substances (PFAS) tables, nhanes 2011-2018*. Centers for Disease Control and Prevention. Retrieved March 1, 2022, from https://www.cdc.gov/exposurereport/pfas_early_release.html

Chang PJ, Bhavsar PK, Michaeloudes C, et al. Corticosteroid insensitivity of chemokine expression in airway smooth muscle of patients with severe asthma. J Allergy Clin Immunol 2012;130:877-85.e5.

Chapman K, Holmes M, Seckl J. 11β-hydroxysteroid dehydrogenases: intracellular gate-keepers of tissue glucocorticoid action. Physiol Rev 2013;93:1139-206.

Cordner, A., Vanessa, Y., Schaider, L. A., Rudel, R. A., Richter, L., & Brown, P. (2019). Guideline levels for PFOA and PFOS in drinking water: the role of scientific uncertainty, risk assessment decisions, and social factors. *Journal of exposure science & environmental epidemiology*, *29*(2), 157-171.

Cousins, I. T., DeWitt, J. C., Glüge, J., Goldenman, G., Herzke, D., Lohmann, R., ... & Wang, Z. (2020). Strategies for grouping per-and polyfluoroalkyl substances (PFAS) to protect human and environmental health. *Environmental Science: Processes & Impacts*, *22*(7), 1444-1460.

Dong GH, Liu MM, Wang D, Zheng L, Liang ZF, Jin YH. 2011. Sub-chronic effect of perfluorooctanesulfonate (PFOS) on the balance of type 1 and type 2 cytokine in adult C57BL6 mice. Arch Toxicol 85:1235–1244.

Dong, G.H.; Tung, K.Y.; Tsai, C.H.; Liu, M.M.; Wang, D.; Liu,W.; Jin, Y.H.; Hsieh,W.S.; Lee, Y.L.; Chen, P.C. Serum polyfluoroalkyl concentrations, asthma outcomes, and immunological markers in a case-control study of Taiwanese children. Environ. Health Perspect. 2013, 121, 507–513.

Eberly, L. E. (2007). Multiple linear regression. *Topics in Biostatistics*, 165-187.

Ege, M. J., Mayer, M., Normand, A. C., Genuneit, J., Cookson, W. O., Braun-Fahrländer, C., ... & von Mutius, E. (2011). Exposure to environmental microorganisms and childhood asthma. *New England Journal of Medicine*, *364*(8), 701-709.

Environmental Protection Agency. (2021). *PFAS Explained*. EPA. Retrieved December 5, 2021. from https://www.epa.gov/pfas/basic-information-pfas

Fairley KJ, Purdy R, Kearns S, Anderson SE, Meade BJ. 2007. Exposure to the immunosuppressant, perfluorooctanoic acid, enhances the murine IgE and airway hyperreactivity response to ovalbumin. Toxicol Sci 97:375–383.

Fuseini, H., Newcomb, D.C., 2017. Mechanisms driving gender differences in asthma. Curr. Allergy Asthma Rep. 17 (3), 19-19.

Graber, J. M., Alexander, C., Laumbach, R. J., Black, K., Strickland, P. O., Georgopoulos, P. G., ... & Weisel, C. P. (2019). Per and polyfluoroalkyl substances (PFAS) blood levels after contamination of a community water supply and comparison with 2013–2014 NHANES. *Journal of exposure science & environmental epidemiology*, *29*(2), 172-182.

Guilbert, T., Zeiger, R. S., Haselkorn, T., Iqbal, A., Alvarez, C., Mink, D. R., ... & Szefler, S. J. (2019). Racial disparities in asthma-related health outcomes in children with severe/difficult-to-treat asthma. *The Journal of Allergy and Clinical Immunology: In Practice*, *7*(2), 568-577.

Hoaglin, D. C. (2003). John W. Tukey and data analysis. *Statistical Science*, 311-318.

Humblet, O.; Diaz-Ramirez, L.G.; Balmes, J.R.; Pinney, S.M.; Hiatt, R.A. Perfluoroalkyl chemicals and asthma among children 12–19 years of age: NHANES (1999–2008). Environ. Health Perspect. 2014, 122, 1129–1133.

Jackson-Browne MS, Eliot M, Patti M, Spanier AJ, Braun JM. PFAS (per- and polyfluoroalkyl substances) and asthma in young children: NHANES 2013-2014. Int J Hyg Environ Health. 2020 Aug;229:113565.

Lee, E. G., Boone, D. L., Chai, S., Libby, S. L., Chien, M., Lodolce, J. P., & Ma, A. (2000). Failure to regulate TNF-induced NF-κB and cell death responses in A20-deficient mice. *Science*, *289*(5488), 2350-2354.

Loewen M, Basu S, Halayko AJ, Kozyrskyj A, Bondy G, Becker AB. 2011. The impact of a perfluorinated compound (PFC) on airway function in an allergic murine model [Abstract]. Am J Respir Crit Care Med 183:A3249.

Miles, J. (2005). R-squared, adjusted R-squared. *Encyclopedia of statistics in behavioral science*.

Myers, R. H (1990). *Classical and modern regression with applications* (Vol. 2, p. 488). Belmont, CA: Duxbury press.

Nardone, A., Rudolph, K. E., Morello-Frosch, R., & Casey, J. A. (2021). Redlines and greenspace: the relationship between historical redlining and 2010 greenspace across the United States. *Environmental health perspectives*, *129*(1), 017006.

OEHHA. (2021). *Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonic Acid (PFOS) in Drinking Water*. Oehha.ca.gov. Retrieved March 1, 2022, from https://oehha.ca.gov/water/report/perfluorooctanoic-acid-pfoa-and-perfluorooctane-sulfonic-acid-pfos-drinking-water

Park SK, Peng Q, Ding N, Mukherjee B, Harlow SD. Determinants of per- and polyfluoroalkyl substances (PFAS) in midlife women: Evidence of racial/ethnic and geographic differences in PFAS exposure. Environ Res. 2019 Aug;175:186-199.

Perperoglou, A., Sauerbrei, W., Abrahamowicz, M. *et al.* A review of spline function procedures in R. *BMC Med Res Methodol* 19, 46 (2019).

Pontius, F. (2019). Regulation of perfluorooctanoic acid (pfoa) and perfluorooctane sulfonic acid (pfos) in drinking water: A comprehensive review. *Water*, *11*(10), 2003.
Post, G. B. (2021). Recent US state and federal drinking water guidelines for per-and polyfluoroalkyl substances. *Environmental Toxicology and Chemistry*, *40*(3), 550-563.

Qin, X.D.; Qian, Z.M.; Dharmage, S.C.; Perret, J.; Geiger, S.D.; Rigdon, S.E.; Howard, S.; Zeng, X.W.; Hu, L.W.; Yang, B.Y.; et al. Association of perfluoroalkyl substances exposure with impaired lung function in children. Environ. Res. 2017, 155, 15–21.

Riedler, J., Braun-Fahrländer, C., Eder, W., Schreuer, M., Waser, M., Maisch, S., ... & ALEX Study Team. (2001). Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. *The Lancet*, *358*(9288), 1129-1133.

Rivera Rivera, N.Y., Tamayo-Ortiz, M., Mercado García, A., Just, A.C., Kloog, I., Tellez-Rojo, M.M., Rosa, M.J., 2020. Prenatal and early life exposure to particulate matter, environmental tobacco smoke and respiratory symptoms in Mexican children. Environ. Res. 110365.

Rosinger, A. Y., Herrick, K. A., Wutich, A. Y., Yoder, J. S., & Ogden, C. L. (2018). Disparities in plain, tap and bottled water consumption among US adults: National Health and Nutrition Examination Survey (NHANES) 2007–2014. *Public health nutrition*, *21*(8), 1455-1464.

Schuijs, M. J., Willart, M. A., Vergote, K., Gras, D., Deswarte, K., Ege, M. J., ... & Hammad, H. (2015). Farm dust and endotoxin protect against allergy through A20 induction in lung epithelial cells. *Science*, *349*(6252), 1106-1110.

Stapleton, M., Howard-Thompson, A., George, C., Hoover, R. M., & Self, T. H. (2011). Smoking and asthma. *The Journal of the American Board of Family Medicine*, *24*(3), 313-322.

Stein, C.R.; McGovern, K.J.; Pajak, A.M.; Maglione, P.J.; Wolff, M.S. Perfluoroalkyl and polyfluoroalkyl substances and indicators of immune function in children aged 12–19 years: National Health and Nutrition Examination Survey. Pediatr. Res. 2016, 79, 348–357.

Stern, L., Berman, J., Lumry, W., Katz, L., Wang, L., Rosenblatt, L., & Doyle, J. J. (2006). Medication compliance and disease exacerbation in patients with asthma: a retrospective study of managed care data. *Annals of Allergy, Asthma & Immunology*, *97*(3), 402-408.

Thomas, A. J., Eberly, L. E., Davey Smith, G., & Neaton, J. D. (2006). ZIP-code-based versus tract-based income measures as long-term risk-adjusted mortality predictors. *American journal of epidemiology*, *164*(6), 586-590.

Tracking California, Public Health Institute. Asthma Related Emergency Department & Hospitalization data. Accessed April 13, 2021 from www.trackingcalifornia.org/asthma/query

Von Mutius, E., & Vercelli, D. (2010). Farm living: effects on childhood asthma and allergy. *Nature Reviews Immunology*, *10*(12), 861-868.

Wang M, Li Q, Hou M, Chan LLY, Liu M, Ter SK, Dong T, Xia Y, Chotirmall SH, Fang M. Inactivation of common airborne antigens by perfluoroalkyl chemicals modulates early life allergic asthma. Proc Natl Acad Sci U S A. 2021 Jun 15;118(24):e2011957118.

Ye L, Zhao B, Cai XH, et al. The inhibitory effects of perfluoroalkyl substances on human and rat 11β-hydroxysteroid dehydrogenase 1. Chem Biol Interact 2012;195:114-8.

Zeng Z, Ma W, Zhao R, Dong X. Airway exposure to perfluorooctanoate exacerbates airway hyperresponsiveness and downregulates glucocorticoid receptor expression in asthmatic mice. Transl Pediatr. 2021 Feb;10(2):323-332.

Zhou J , Baumann K , Mead RN , Skrabal SA , Kieber RJ , Avery GB , Shimizu M , DeWitt JC , Sun M , Vance SA , Bodnar W , Zhang Z , Collins LB , Surratt JD , Turpin BJ . PFOS dominates PFAS composition in ambient fine particulate matter (PM2.5) collected across North Carolina nearly 20 years after the end of its US production. Environ Sci Process Impacts. 2021 Apr 1;23(4):580-587. doi: 10.1039/d0em00497a. Epub 2021 Mar 16. PMID: 33725038.

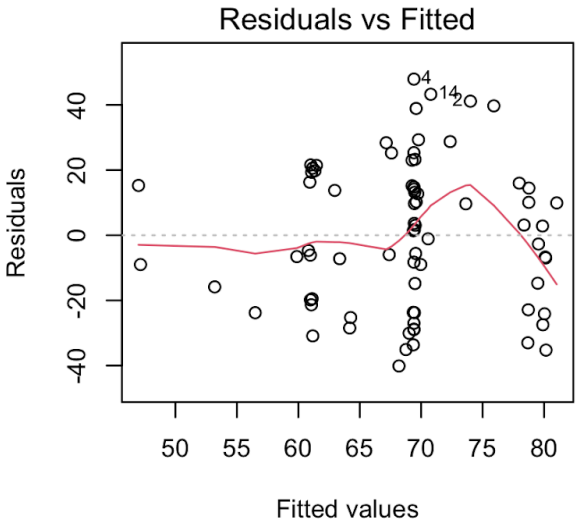
Zhu, Y.; Qin, X.D.; Zeng, X.W.; Paul, G.; Morawska, L.; Su, M.W.; Tsai, C.H.;Wang, S.Q.; Lee, Y.L.; Dong, G.H. Associations of serum perfluoroalkyl acid levels with T-helper cell-specific cytokines in children: By gender and asthma status. Sci. Total Environ. 2016, 559, 166–173.

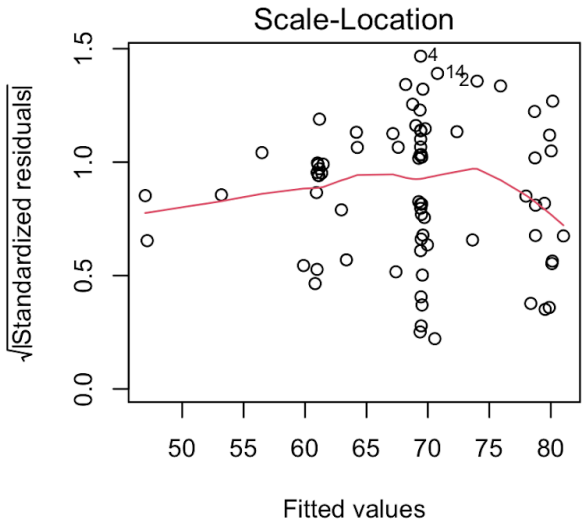
Appendix A: R Code

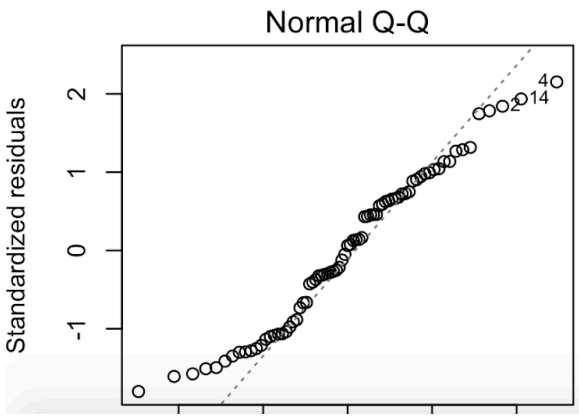
```
> library(readr)
> pfos_pfoa <- read_csv("~/downloads/PFAS_Asthma - Background Weighted - Target Data [wt].csv")
cols(
  Zip = col_double(),
  `PFOS [wt]` = col_double(),
  `PFOA[wt]` = col_double(),
  `Asthma ED [Avg]` = col_double()
)
> summary(pfos_pfoa)
      Zip          PFOS [wt]         PFOA[wt]     Asthma ED [Avg]
 Min.   :90022   Min.   : 5.000   Min.   :2.000   Min.   : 28.07
 1st Qu.:90634   1st Qu.: 5.011   1st Qu.:2.000   1st Qu.: 45.67
 Median :91382   Median : 5.606   Median :2.003   Median : 72.57
 Mean   :91639   Mean   : 7.381   Mean   :2.557   Mean   : 68.86
 3rd Qu.:92822   3rd Qu.:10.093   3rd Qu.:3.130   3rd Qu.: 83.40
 Max.   :92887   Max.   :11.299   Max.   :4.448   Max.   :117.33

#Histograms & Boxplots
> hist(pfos_pfoa$'PFOS [wt]')
> hist(pfos_pfoa$'PFOA[wt]')
> library(ggplot2)
> ggplot(data = pfos_pfoa,
+        aes(x = `PFOS [wt]`, y =  `Asthma ED [Avg]`)) +
+   geom_point()
> ggplot(data = pfos_pfoa,
+        aes(x = `PFOA [wt]`, y =  `Asthma ED [Avg]`)) +
+   geom_point()

#exploratory analysis
> pfos_lm_fit <- lm(`Asthma ED [Avg]`~`PFOS [wt]`, data = pfos_pfoa)
> summary(pfos_lm_fit)

Call:
lm(formula = `Asthma ED [Avg]` ~ `PFOS [wt]`, data = pfos_pfoa)

Residuals:
```

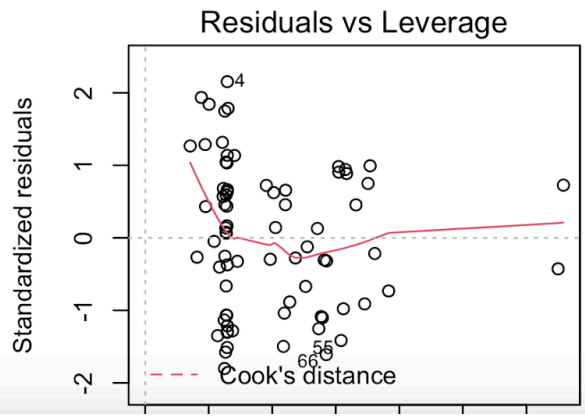
```
   Min      1Q  Median      3Q     Max
-40.673 -21.757   3.341  15.675  49.924

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  64.3703     8.2880   7.767 3.86e-11 ***
`PFOS [wt]`   0.6077     1.0607   0.573    0.568
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23.56 on 73 degrees of freedom
Multiple R-squared:  0.004476,    Adjusted R-squared:  -0.009161
F-statistic: 0.3282 on 1 and 73 DF,  p-value: 0.5685

> pfoa_lm_fit <- lm(`Asthma ED [Avg]`~`PFOA[wt]`, data = pfos_pfoa)
> summary(pfoa_lm_fit)

Call:
lm(formula = `Asthma ED [Avg]` ~ `PFOA[wt]`, data = pfos_pfoa)

Residuals:
   Min      1Q  Median      3Q     Max
-43.664 -18.109   0.471  19.711  44.704

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   86.186      8.203  10.507 2.97e-16 ***
`PFOA[wt]`    -6.778      3.038  -2.231   0.0287 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22.85 on 73 degrees of freedom
Multiple R-squared:  0.06384,  Adjusted R-squared:  0.05102
F-statistic: 4.978 on 1 and 73 DF,  p-value: 0.02874

> combined_lm_fit <- lm(`Asthma ED [Avg]`~`PFOS [wt]`+`PFOA[wt]`, data = pfos_pfoa)
> summary(combined_lm_fit)

Call:
lm(formula = `Asthma ED [Avg]` ~ `PFOS [wt]` + `PFOA[wt]`, data = pfos_pfoa)
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-40.123 -20.591   1.404  15.626  47.930

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   78.152      9.366   8.344  3.5e-12 ***
`PFOS [wt]`    1.910      1.119   1.707  0.09219 .
`PFOA[wt]`    -9.149      3.305  -2.768  0.00716 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 22.55 on 72 degrees of freedom
Multiple R-squared:  0.1002,    Adjusted R-squared:  0.07525
F-statistic: 4.011 on 2 and 72 DF,  p-value: 0.02231

# diagnostic plots
> layout(matrix(c(1,2,3,4),2,2))
> plot(combined_lm_fit)
```










```
> anova(combined_lm_fit,pfoa_lm_fit)
Analysis of Variance Table

Model 1: `Asthma ED [Avg]` ~ `PFOS [wt]` + `PFOA[wt]`
Model 2: `Asthma ED [Avg]` ~ `PFOA[wt]`
  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1     72 36624
2     73 38105 -1   -1481.7 2.9129 0.09219 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

#Regression Analysis

```
> model<-lm(df$'Asthma ED [Avg]' ~ df$'PFOS [wt]' + df$'PFOA[wt]'+df$pm2.5 + df$diesel+df$income_qrt + df$race_white +
df$edu, data = df)
> summary(model)

Call:
lm(formula = df$"Asthma ED [Avg]" ~ df$"PFOS [wt]" + df$"PFOA[wt]" +
    df$pm2.5 + df$diesel + df$income_qrt + df$race_white + df$edu,
    data = df)

Residuals:
    Min      1Q  Median      3Q     Max
-38.540 -11.425  -3.092   8.366  51.451

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    115.6986    12.1134   9.551 3.98e-14 ***
df$"PFOS [wt]"   0.3987     0.9035   0.441   0.6604
df$"PFOA[wt]"   -0.3655     2.8947  -0.126   0.8999
df$pm2.5        -1.2999     0.7686  -1.691   0.0954 .
df$diesel       36.2314    25.6952   1.410   0.1632
df$income_qrt  -13.7728     2.9499  -4.669 1.50e-05 ***
df$race_white  -18.7722    15.6600  -1.199   0.2349
df$edu          -4.6905    46.3049  -0.101   0.9196
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17 on 67 degrees of freedom
Multiple R-squared:  0.5246,    Adjusted R-squared:  0.4749
F-statistic: 10.56 on 7 and 67 DF,  p-value: 7.187e-09
```

```
Call:
lm(formula = df.f$"Asthma ED [Avg]" ~ PFOS_Q + PFOA_Q + pm2.5_Q +
    diesel_Q + income_Q + race_Q + edu_Q, data = df.f)

Residuals:
    Min      1Q  Median      3Q     Max
-32.620 -11.310  -0.113   7.859  32.127

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  79.4281    10.1103   7.856  1.9e-10 ***
PFOS_Q2       0.3151     6.1666   0.051 0.959435
PFOS_Q3      10.1112     7.1357   1.417 0.162339
PFOS_Q4      -2.1977     7.2345  -0.304 0.762483
PFOA_Q2       0.7310     6.2693   0.117 0.907623
PFOA_Q3       0.1114     6.4963   0.017 0.986383
PFOA_Q4      -1.7896     7.4426  -0.240 0.810907
pm2.5_Q2     -3.3852     9.7651  -0.347 0.730215
pm2.5_Q3     -4.0394    11.2604  -0.359 0.721224
pm2.5_Q4    -12.0968    11.6727  -1.036 0.304756
diesel_Q2     9.1447    10.4693   0.873 0.386340
diesel_Q3    10.2981    11.9422   0.862 0.392390
diesel_Q4    10.1518    13.1314   0.773 0.442901
income_Q2    -4.4672     6.0681  -0.736 0.464872
income_Q3   -16.2827     7.9039  -2.060 0.044315 *
income_Q4   -31.6931     7.9795  -3.972 0.000216 ***
race_Q2       4.1486     6.8460   0.606 0.547105
race_Q3       1.1203     6.8833   0.163 0.871335
race_Q4     -11.2043     8.3644  -1.340 0.186118
edu_Q2       -0.7153     6.4081  -0.112 0.911540
edu_Q3       -7.5289     5.6500  -1.333 0.188379
edu_Q4        5.2854     6.1743   0.856 0.395834
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.43 on 53 degrees of freedom
```

Multiple R-squared: 0.6486, Adjusted R-squared: 0.5093

F-statistic: 4.658 on 21 and 53 DF, p-value: 2.901e-06

```
> df.cor = cor(df, method = c("spearman"))
> as.matrix(df.cor)
              Asthma ED [Avg]      PFOS_Q      PFOA_Q     pm2.5_Q    diesel_Q   income_Q      race_Q       edu_Q
Asthma ED [Avg]       1.0000000  0.19808216 -0.33385872  0.11894038  0.28781633 -0.6759445 -0.48680206  0.19734781
PFOS_Q                0.1980822  1.00000000  0.20398719  0.30893933  0.28618137 -0.2290058 -0.01646167  0.02668083
PFOA_Q               -0.3338587  0.20398719  1.00000000 -0.04594832 -0.16500281  0.4386673  0.49511986 -0.15283120
pm2.5_Q               0.1189404  0.30893933 -0.04594832  1.00000000  0.81491709 -0.1647070 -0.18925217  0.03733339
diesel_Q              0.2878163  0.28618137 -0.16500281  0.81491709  1.00000000 -0.3708182 -0.34468378  0.04857143
income_Q             -0.6759445 -0.22900578  0.43866729 -0.16470695 -0.37081822  1.0000000  0.55230610 -0.36907594
race_Q               -0.4868021 -0.01646167  0.49511986 -0.18925217 -0.34468378  0.5523061  1.00000000 -0.13152560
edu_Q                 0.1973478  0.02668083 -0.15283120  0.03733339  0.04857143 -0.3690759 -0.13152560  1.00000000

> model1<-lm(formula=df.f$'Asthma ED [Avg]'~PFOS_Q+PFOA_Q+diesel_Q+race_Q+edu_Q, data=df.f)
> summary(model1)

Call:
lm(formula = df.f$"Asthma ED [Avg]" ~ PFOS_Q + PFOA_Q + diesel_Q +
    race_Q + edu_Q, data = df.f)

Residuals:
    Min      1Q  Median      3Q     Max
-34.304 -11.576   0.061  14.419  31.637

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  70.5910    10.3212   6.839 5.06e-09 ***
PFOS_Q2       5.8234     6.7055   0.868  0.38867
PFOS_Q3      14.5515     7.1315   2.040  0.04579 *
PFOS_Q4       7.6505     7.1706   1.067  0.29035
PFOA_Q2      -4.1262     6.7124  -0.615  0.54111
PFOA_Q3      -7.3344     6.9373  -1.057  0.29471
PFOA_Q4      -9.3192     8.2254  -1.133  0.26181
diesel_Q2     5.0265     6.8983   0.729  0.46909
```

```
diesel_Q3     5.4467     7.9081   0.689  0.49369
diesel_Q4     3.3504     7.8167   0.429  0.66976
race_Q2      -0.8806     6.9331  -0.127  0.89936
race_Q3      -8.9480     7.2070  -1.242  0.21930
race_Q4     -26.2584     8.2614  -3.178  0.00236 **
edu_Q2        0.4421     7.1372   0.062  0.95082
edu_Q3       -8.2529     6.4049  -1.289  0.20259
edu_Q4       13.0686     6.3942   2.044  0.04544 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18.87 on 59 degrees of freedom
Multiple R-squared:  0.4839,    Adjusted R-squared:  0.3527
F-statistic: 3.688 on 15 and 59 DF,  p-value: 0.0001575


> model2<-lm(formula=df.f$'Asthma ED [Avg]'~PFOS_Q+PFOA_Q+diesel_Q+income_Q+edu_Q, data=df.f)
> summary(model2)

Call:
lm(formula = df.f$"Asthma ED [Avg]" ~ PFOS_Q + PFOA_Q + diesel_Q +
    income_Q + edu_Q, data = df.f)

Residuals:
    Min      1Q  Median      3Q     Max
-30.555 -11.768   0.300   9.075  35.813

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  79.1911     9.1357   8.668 4.10e-12 ***
PFOS_Q2       0.9354     6.0946   0.153   0.8785
PFOS_Q3      10.6301     6.4801   1.640   0.1062
PFOS_Q4      -1.2689     6.8740  -0.185   0.8542
PFOA_Q2       1.1023     5.9202   0.186   0.8529
PFOA_Q3       1.6473     6.4720   0.255   0.8000
PFOA_Q4      -3.7004     7.2241  -0.512   0.6104
diesel_Q2     6.5533     6.2866   1.042   0.3015
diesel_Q3     2.3298     6.6939   0.348   0.7290
diesel_Q4     2.0427     6.5880   0.310   0.7576
```

```
income_Q2    -3.2320     6.0014  -0.539   0.5922
income_Q3   -18.7888     7.3162  -2.568   0.0128 *
income_Q4   -37.4295     7.2658  -5.151 3.13e-06 ***
edu_Q2       -1.0129     6.3155  -0.160   0.8731
edu_Q3       -5.8007     5.4444  -1.065   0.2910
edu_Q4        5.2576     6.0346   0.871   0.3872
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.57 on 59 degrees of freedom
Multiple R-squared:  0.6022,    Adjusted R-squared:  0.5011
F-statistic: 5.954 on 15 and 59 DF,  p-value: 2.738e-07


> model3<-lm(formula=df.f$'Asthma ED [Avg]'~PFOS_Q+PFOA_Q+pm2.5_Q+income_Q+edu_Q, data=df.f)
> summary(model3)

Call:
lm(formula = df.f$"Asthma ED [Avg]" ~ PFOS_Q + PFOA_Q + pm2.5_Q +
    income_Q + edu_Q, data = df.f)

Residuals:
    Min      1Q  Median      3Q     Max
-28.963 -11.450  -0.105   7.785  38.248

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  79.27694    8.95277   8.855 1.99e-12 ***
PFOS_Q2       2.01689    5.95870   0.338   0.7362
PFOS_Q3      13.48162    6.42839   2.097   0.0403 *
PFOS_Q4      -0.09775    6.81971  -0.014   0.9886
PFOA_Q2       2.05265    6.02710   0.341   0.7346
PFOA_Q3      -0.04821    6.29419  -0.008   0.9939
PFOA_Q4      -4.22185    7.00967  -0.602   0.5493
pm2.5_Q2      3.83540    5.90443   0.650   0.5185
pm2.5_Q3      2.53589    6.14400   0.413   0.6813
pm2.5_Q4     -3.40058    6.07789  -0.559   0.5779
income_Q2    -2.32945    5.87316  -0.397   0.6931
income_Q3   -17.22771    7.06232  -2.439   0.0177 *
```

```
income_Q4   -35.70170    7.03290  -5.076 4.12e-06 ***
edu_Q2       -2.23347    6.12613  -0.365   0.7167
edu_Q3       -5.88103    5.42586  -1.084   0.2828
edu_Q4        5.92384    5.97406   0.992   0.3254
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.51 on 59 degrees of freedom
Multiple R-squared:  0.605,      Adjusted R-squared:  0.5046
F-statistic: 6.024 on 15 and 59 DF,  p-value: 2.287e-07
> model4<-lm(formula=df.f$'Asthma ED [Avg]'~PFOS_Q+PFOA_Q+pm2.5_Q+race_Q+edu_Q, data=df.f)
> summary(model4)

Call:
lm(formula = df.f$"Asthma ED [Avg]" ~ PFOS_Q + PFOA_Q + pm2.5_Q +
    race_Q + edu_Q, data = df.f)

Residuals:
    Min      1Q  Median      3Q     Max
-37.001 -11.568   0.251  14.253  38.289

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  71.4170     9.8396   7.258 9.89e-10 ***
PFOS_Q2       6.2122     6.4528   0.963  0.33962
PFOS_Q3      18.0556     6.8454   2.638  0.01066 *
PFOS_Q4       7.6670     7.0160   1.093  0.27893
PFOA_Q2      -1.3345     6.7232  -0.198  0.84334
PFOA_Q3      -9.0383     6.5577  -1.378  0.17333
PFOA_Q4      -7.0505     7.8370  -0.900  0.37197
pm2.5_Q2      5.1864     6.3695   0.814  0.41877
pm2.5_Q3      4.2314     6.9980   0.605  0.54772
pm2.5_Q4     -6.9442     6.9965  -0.993  0.32499
race_Q2       0.8312     6.7452   0.123  0.90235
race_Q3     -10.1878     6.7236  -1.515  0.13505
race_Q4     -26.4512     7.8577  -3.366  0.00135 **
edu_Q2        0.2296     6.8084   0.034  0.97321
edu_Q3       -7.8667     6.2058  -1.268  0.20991
```

```
edu_Q4       13.7404     6.1288   2.242  0.02874 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18.36 on 59 degrees of freedom
Multiple R-squared:  0.5112,    Adjusted R-squared:  0.3869
F-statistic: 4.113 on 15 and 59 DF,  p-value: 4.422e-05


> model_fs_spline<-gam(df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs="fs") + s(PFOA_Q, bs="fs") + s(pm2.5_Q, bs="fs") +
+     s(income_Q, bs="fs") + s(edu_Q, bs="fs"), , data=df.f)
> summary.gam(model_fs_spline)

Family: gaussian
Link function: identity

Formula:
df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs = "fs") + s(PFOA_Q, bs = "fs") +
    s(pm2.5_Q, bs = "fs") + s(income_Q, bs = "fs") + s(edu_Q,
    bs = "fs")

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   69.051      8.071   8.555 2.02e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df      F p-value
s(PFOS_Q)   1.710e+00      3  1.938   0.142
s(PFOA_Q)   3.371e-11      4  0.000   0.761
s(pm2.5_Q)  3.581e-11      4  0.000   0.820
s(income_Q) 2.811e+00      3 28.876  <2e-16 ***
s(edu_Q)    8.972e-01      3  0.702   0.198
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.533   Deviance explained = 56.7%
GCV = 281.19  Scale est. = 257.13    n = 75
```

```
> model_fs2_spline<-gam(df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs="fs") + s(PFOA_Q, bs="fs") + s(pm2.5_Q, bs="fs") +
+     s(race_Q, bs="fs") + s(edu_Q, bs="fs"), , data=df.f)
> summary.gam(model_fs2_spline)

Family: gaussian
Link function: identity

Formula:
df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs = "fs") + s(PFOA_Q, bs = "fs") +
    s(pm2.5_Q, bs = "fs") + s(race_Q, bs = "fs") + s(edu_Q, bs = "fs")

Parametric coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   68.330      8.322   8.211 9.64e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df      F  p-value
s(PFOS_Q)  1.607e+00      3  1.844   0.0621 .
s(PFOA_Q)  5.585e-11      4  0.000   0.4792
s(pm2.5_Q) 1.035e-08      3  0.000   0.4288
s(race_Q)  2.697e+00      3 11.209 1.81e-05 ***
s(edu_Q)   2.294e+00      3  4.342   0.0106 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.39   Deviance explained = 44.4%
GCV =  373.6  Scale est. = 335.75    n = 75
> model_fs3_spline<-gam(df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs="fs") + s(PFOA_Q, bs="fs") + s(diesel_Q, bs="fs") +
+     s(race_Q, bs="fs") + s(edu_Q, bs="fs"), , data=df.f)
> summary.gam(model_fs3_spline)

Family: gaussian
Link function: identity

Formula:
df.f$"Asthma ED [Avg]" ~ s(PFOS_Q, bs = "fs") + s(PFOA_Q, bs = "fs") +
    s(diesel_Q, bs = "fs") + s(race_Q, bs = "fs") + s(edu_Q,
```

```
   bs = "fs")

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   68.330      8.322   8.211 9.64e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df      F  p-value
s(PFOS_Q)   1.607e+00      3  1.844   0.0621 .
s(PFOA_Q)   9.387e-11      4  0.000   0.4797
s(diesel_Q) 8.081e-11      4  0.000   0.5447
s(race_Q)   2.697e+00      3 11.209 1.81e-05 ***
s(edu_Q)    2.294e+00      3  4.342   0.0106 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.39   Deviance explained = 44.4%
GCV =  373.6  Scale est. = 335.75    n = 75

---
> model_dfs2_spline<-gam(df.s$"Asthma ED [Avg]" ~ s(PFOS_W) + s(PFOA_W) + pm2.5_Q +
+     income_Q+ edu_Q, , data=df.s)
> summary(model_dfs2_spline)

Family: gaussian
Link function: identity

Formula:
df.s$"Asthma ED [Avg]" ~ s(PFOS_W) + s(PFOA_W) + pm2.5_Q + income_Q +
   edu_Q

Parametric coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  81.407373   6.767117  12.030  < 2e-16 ***
pm2.5_Q2      4.606608   5.678782   0.811  0.42042
pm2.5_Q3      1.006703   5.783547   0.174  0.86239
pm2.5_Q4     -0.005333   5.783079  -0.001  0.99927
```

```
income_Q2     0.850215   5.767805   0.147  0.88330
income_Q3   -17.317986   6.285953  -2.755  0.00773 **
income_Q4   -34.524862   6.257932  -5.517 7.52e-07 ***
edu_Q2       -1.658857   5.587766  -0.297  0.76758
edu_Q3       -6.556201   5.096436  -1.286  0.20317
edu_Q4        3.909590   5.709913   0.685  0.49614
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
           edf Ref.df     F p-value
s(PFOS_W) 3.191  3.921 2.147  0.0759 .
s(PFOA_W) 1.062  1.120 0.528  0.4470
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.541   Deviance explained = 62.3%
GCV = 311.97  Scale est. = 252.68     n = 75



#spline plot
> grdpts = seq(min(df.s$PFOS_W),max(df.s$PFOS_W),length.out=1000)
> head(grdpts)
[1] 5.000000 5.006305 5.012610 5.018916 5.025221 5.031526
> grdpots
> prdgrd = data.frame(PFOS_W=grdpts, pm2.5_W=median(df.s$pm2.5_W,na.rm=T),PFOA_W=median(df.s$PFOA_W,na.rm=T),
income_Q=2, edu_Q=2)

>  plot(grdpts, type="l", xlab="PFOS", ylab = "Asthma ED",
+
+  main="pfos and asthma ed")
```