Air Quality and Diabetes in California

Hope Cupples, Angie Bouche, Vanessa Guenter, Yirui Zhang
Faculty Advisor: Ashley Larsen PhD Advisor: Patrick Hunnicutt

James S. Bower Foundation Re-envisioning. | Sansum Diabetes Research Institute | Bren School of Environmental Science & Management, University of California, Santa Barbara

Background and Significance

Mil Familias, a project of Sansum Diabetes Research Institute

Sansum Diabetes Research Institute (SDRI) has a 75-year tenure conducting research, education, and care in diabetes. Incited by the disproportionate burden of diabetes on Latino families and recognizing that effective diabetes prevention and treatment rest on understanding the determinants of diabetes as well as accurate risk calculation, SDRI launched the Mil Familias project (www.milfamilias.com) in 2018.

California is home to 6 of the 10 most air polluted cities in the US[1]

Diabetes costs California more than $27 billion annually [2]

Diabetes has a higher prevalence among minorities [3]

A specific pollutant, particulate matter ($PM_{2.5}$) has been linked to a range of negative health impacts including diabetes. There has been no study exploring this association in California. This project aims to fill this crucial gap by assessing relationships between particulate matter 2.5 ($PM_{2.5}$) and diabetes prevalence throughout the state using a cross sectional and panel data approach.

Research Questions

1. **What are the yearly average $PM_{2.5}$ concentrations for each census tract in California?**
2. **What is the relationship between $PM_{2.5}$ concentration and diabetes prevalence in California?**
3. **How do these relationships change if we incorporate threshold values for $PM_{2.5}$ levels or demographic subgroups?**

[1]American Lung Association. 2017. State of the Air, Most Polluted Cities.
[2]American Diabetes Association 2018. Facts About Type 2.
[3]Office of Minority Health. 2016. Diabetes and Hispanic Americans.

What is Diabetes?

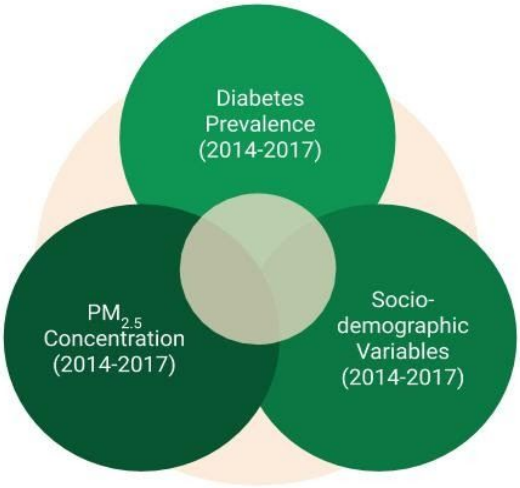
Diabetes occurs when the pancreas does not produce enough insulin (type 1), or when the body cannot effectively use the insulin it produces (type 2). Type 2 diabetes represents 95% of all diabetes cases.

Diabetes of either type can lead to:

- Blindness
- Stroke
- Heart attack
- Death

Type 1 diabetes is hereditary, but type 2 is preventable. Well established risk factors contributing to type 2 diabetes include family history, sedentary lifestyle, unhealthy diet, and age.

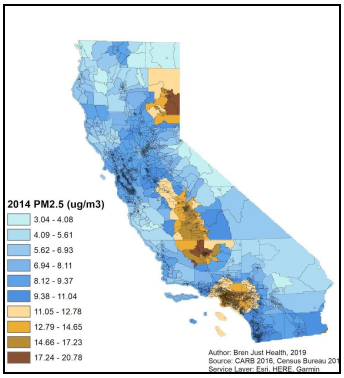
1 Data

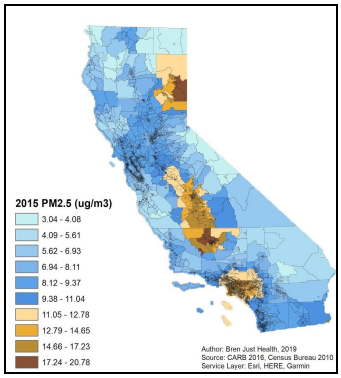


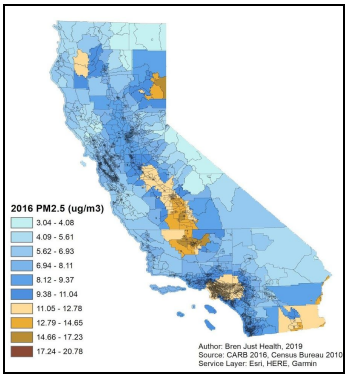
In this project we combine publically available datasets for the years 2014-2017:

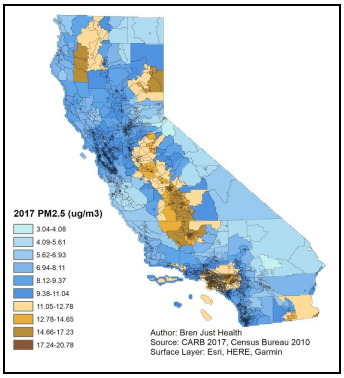
1) $PM_{2.5}$ concentrations from the California Air Resource Board
2) Diabetes prevalence from the Centers of Disease Control
3) Sociodemographic variables from the Census Bureau's American Community Survey.

The $PM_{2.5}$ data was recorded as daily or hourly measurements from air quality monitors throughout the state, but for our statistical tests we needed to convert these values to an average annual $PM_{2.5}$ value for each census tract. Diabetes prevalence and socioeconomic variables were already provided in this format. We first averaged annual $PM_{2.5}$ concentrations for each monitoring location and interpolated these concentrations across the state. Annual $PM_{2.5}$ levels for each census tract are shown in the maps below. After wrangling all datasets, we retained observations for 5,084 census tracts across California.





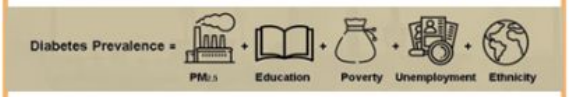




2 Statistical Analysis

We used two model types to assess relationships between $PM_{2.5}$, diabetes prevalence, and sociodemographic indicators. In all models, sociodemographic variables included educational attainment, poverty rate, unemployment rate, and race/ethnicity. Here, $PM_{2.5}$ was calculated as a yearly average.

Cross Sectional Model



Assesses the relationship between diabetes, $PM_{2.5}$, and sociodemographic variables within each year of study.

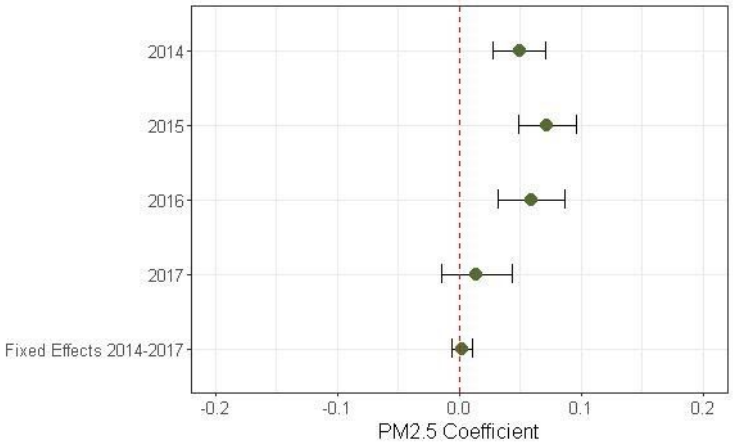
Fixed Effects Model

Uses the same variables as the cross sectional model, but also controls for time-invariant factors between census tracts.

Results

There is a significantly positive association between $PM_{2.5}$ and diabetes prevalence in the years 2014, 2015, and 2016.

In models that include data from 2017, there is no association between $PM_{2.5}$ and diabetes.

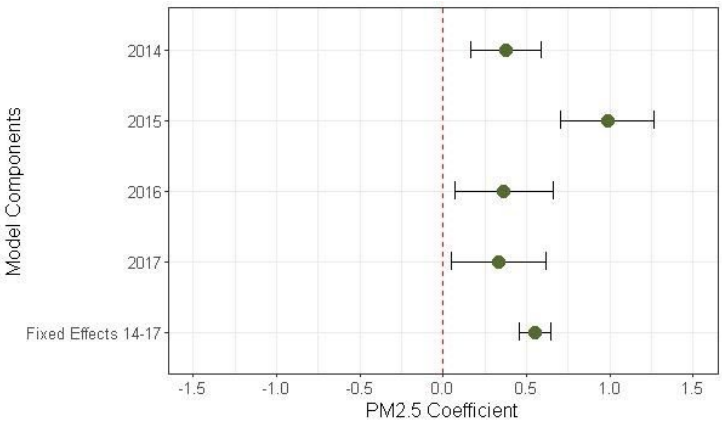


3 Thresholds and Subgroups

We conducted an alternative analysis to explore the possibility of non-linear relationships between $PM_{2.5}$ and diabetes. We calculated $PM_{2.5}$ as the proportion of days in non-attainment of the National Ambient Air Quality Standard (NAAQ), which is 12μg/m^3.

Thresholds

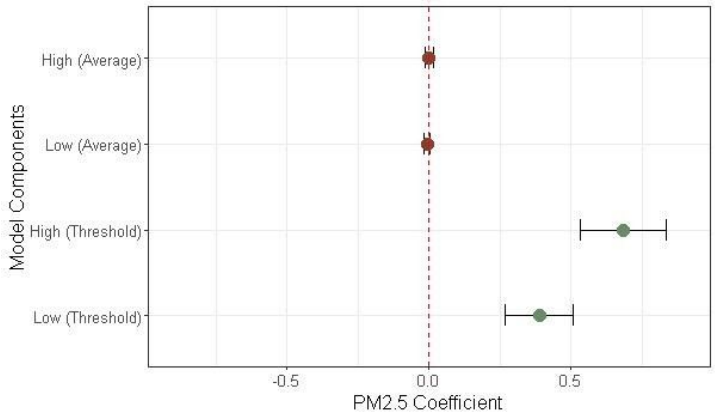
When we incorporate non-attainment days of $PM_{2.5}$ into the models, there is a significant positive relationship ($p<0.001$) between $PM_{2.5}$ and diabetes prevalence across all models and years.



Subgroups

Census tracts in California with a high Latino population (>33% Latino of any race) consistently have higher concentrations of $PM_{2.5}$ and diabetes prevalence compared to census tracts with smaller Latino populations.

	2014		2015		2016		2017	
	Low	High	Low	High	Low	High	Low	High
Diabetes	8.63	11.53	8.22	11.19	8.49	11.15	8.4	11.06
$PM_{2.5}$	10.58	11.90	9.59	10.91	9.48	10.63	10.63	11.53
Unemployment	5.96	8.50	5.33	7.62	4.69	6.69	4.13	5.83
Education	8.84	31.71	8.68	31.17	8.57	30.70	8.38	29.88
Poverty	11.47	23.41	11.36	23.20	11.10	22.48	10.74	21.12



Subgroups

We ran the fixed effects model on both subgroups using $PM_{2.5}$ calculated as yearly averages and as non-attainment days. We see there is a larger positive relationship between $PM_{2.5}$ and diabetes prevalence in heavily Latino-populated census tracts compared to areas with small Latino populations only when we incorporate $PM_{2.5}$ as non-attainment days.

Conclusion

Future Research

The area of public health that incorporates environmental conditions is a new and growing field of research. There are many opportunities for new development, such as the following:

Literature studying the relationships between chronic health conditions and environmental factors is not conclusive on the timeline at which pollution can affect the body.

Diabetes prevalence in the United States differs by race and ethnicity. Additional studies could incorporate interactions between pollution and demographics.

Key Results

1. We find suggestive evidence of a positive association between **average $PM_{2.5}$ concentration** and diabetes prevalence that is sensitive to time and cross-sectional units.

2. We find stronger evidence of an association between **non-attainment days of $PM_{2.5}$** and diabetes prevalence.

3. We see these same trends reflected in Latino subgroups. We see a larger positive association between non-attainment days of $PM_{2.5}$ and diabetes prevalence in high-Latino census tracts compared to low-Latino tracts.