Refinery Pollution and Health: The Relationship Between Air Pollution and Emergency Department Admissions in the San Francisco Bay Area Refinery Corridor

Daniel C. Menza

ABSTRACT

High levels of ambient air pollution have been identified as a risk factor for many acute adverse health outcomes, specifically cardiovascular and respiratory events. This link has not been studied in the context of environmental health disparity, which examines how certain communities and populations, mainly people of color and low socioeconomic status, are exposed to a disproportionately high amount of pollution, and consequently, experience poor health outcomes. Environmental justice advocates have identified the Refinery Corridor in the San Francisco Bay Area as a region where this environmental health disparity exists. The Refinery Corridor is named for the five petroleum refineries in the region, as well as a multitude of chemical manufacturers and other industrial polluters. This study examines the relationship between levels of particulate matter (PM2.5), sulfur dioxide (SO2), and the rate of Emergency Department admissions for cardiovascular and respiratory events in this region. I used a linear regression model to compare the strength of this relationship in the Refinery Corridor and the surrounding communities to investigate whether air pollution impacted health more strongly in the Refinery Corridor. I found that over the whole study area, PM2.5 and SO2 were significantly and positively correlated with respiratory admissions. The relationship was stronger outside the refinery corridor than inside, contrary to my hypothesis. Therefore, the public datasets I utilized for this study were not detailed enough to show this disparity if it does exist. I will repeat this study with more detailed data from the Public Health Department to more conclusively determine if there is an environmental health disparity in the Refinery Corridor.

KEYWORDS

Environmental justice, PM 2.5, sulfur dioxide, environmental health disparity

INTRODUCTION

Pollution is a major environmental determinant of health, and in the United States the health burden of pollution is distributed inequitably. This is especially true in the Refinery Corridor, a region of the San Francisco Bay Area. There are five major petroleum refineries, as well as many other nearby sources of industrial pollution in the Refinery Corridor, which release a large amount of pollution that poses health risks to the surrounding communities (Brody et al 2009; Cohen et al 2012). These Refinery Corridor communities have formed environmental justice (EJ) groups to voice their opposition regarding incurred, disproportionate pollution burdens. These groups often lack the capacity to conduct research that might provide empirical evidence describing this inequitable and dangerous exposure. The goal of this study is to examine the relationship between ambient air pollution and health outcomes in the Refinery Corridor and its surrounding communities, in order to investigate these purported disparities. Results will inform EJ groups and community advocates, along with other invested stakeholders, and guide future research on socio-spatial distributions of health burdens.

Two adverse health outcomes which are known to be highly impacted by pollution, and therefore represent potential mechanisms for this environmental health disparity, are acute cardiovascular and respiratory events. For every 10 $\mu g/m^3$ increase in particulate matter under 2.5 microns in size (PM2.5) the daily risk of cardiovascular mortality increases by 0.4-1% (Brooks et al. 2010). The same increase in PM10 (particulate matter under 10 microns in size) concentration causes a 1% increase in daily mortality from respiratory conditions (Pope and Dockery 1994). Investigators obtained these results by linking daily pollution levels to the rate of hospital admissions for acute cardiovascular and respiratory health outcomes, respectively. This relationship between daily air pollution and hospital visits is well established, and review papers like Brooks et al. (2004; 2010) summarize the literature on this very well. This makes it an attractive framework to use to study environmental health disparity. In order to adapt this study design to this purpose, I will compare the relationship between hospital admissions and pollution levels separately in the Refinery Corridor and surrounding communities, in order investigate environmental health disparity in this area.

Although there has been research done on environmental health in the Refinery Corridor, there are very few studies directly linking daily pollution and acute health in this region. This study examines the relationship between levels of PM 2.5 and Sulfur Dioxide (SO2) and the rates of emergency department visits for cardiovascular and respiratory health events in the Refinery Corridor and surrounding communities. I have chosen these two pollutants because SO2 is a major refinery pollutant that is monitored daily by the EPA. Literature on the link between SO2 and acute health outcomes is not as conclusive as the link established for PM2.5, however, since refinery pollution is so ubiquitous in this area; I believe the pollution-admissions relationship between SO2 and acute health will be strong. Identifying a positive relationship between SO2 and acute health outcomes in the Refinery Corridor would be strong evidence that refinery pollution seriously impacts health. PM2.5 will also be analyzed, although it is not the major pollutant of concern produced from refining. Studying the relationship between air pollution and health outcomes will reveal a way in which the health of Refinery Corridor communities is directly impacted by the refineries and other polluters in the area. Recognizing and understanding the link between refinery pollution and acute health outcomes in this area represents a major step towards understanding how environmental health disparities operate, generally, and identifying the drivers of inequalities in Richmond.

BACKGROUND

The ultimate goal of this study is to determine how the health burden of air pollution varies between the heavily polluted Refinery Corridor and the rest of the San Francisco Bay Area. Understanding this relationship will provide deeper insight into the environmental health disparity that arises from living in an area like the Refinery Corridor, which has a very high concentration of industries like oil refining and chemical manufacturing. In order to do this, I have adapted the design of studies used to relate air pollution to acute health outcomes to visualize how the relationship varies spatially over the Bay Area. This modified design will not only examine the relationship between air pollution and acute health outcomes, but it will also reveal how that relationship is impacted by proximity to large polluters like oil refineries.

Air Pollution-Health Outcomes Studies

Over the last 20 years, a large body of research has shown that high levels of air pollution can acutely increase the risk of acute respiratory and cardiovascular outcomes (Dockery and Pope 1994, Brooks et al. 2010). These studies have largely been conducted by combining hospital admissions data with data from air pollution monitors and using regression models to measure the impact that air pollution has on the daily rates of these health outcomes (Schwartz and Morris 1995; Dominici et al. 2006; Poloniecki et al. 1997). However, to my knowledge, this study design has not previously been applied to issues of environmental health disparity and environmental justice. This design is well suited to study environmental health disparity, because it examines the impact of pollution on health. If pollution has more impact on health in the Refinery Corridor compared to the surrounding areas, that is good evidence of environmental health disparity. In order to adapt this study design to study environmental health disparity, specifically inside and outside of the Refinery Corridor, I will relate the pollutants and admissions separately in the Refinery Corridor and the rest of Contra Costa and Solano county.

In order to gain the most detailed insight into the environmental health disparity in this area and how it is driven by air pollution, detailed daily data on the incidence of the health outcomes of interest would be ideal. However, this data must be requested from the Department of Public Health, which can be a lengthy process. I am currently in the process of obtaining this detailed daily data, and while I wait I will analyze the available public data. The best available public data on health outcomes for this study is the Office of Statewide Health and Planning Department's (OSHPD) Emergency Department Use Reports. These are quarterly reports published by each ED which describe the number of patients seen each period, and groups the admissions by what type of health issue they had. These are far from the daily time scale the protected data will have, but it is the best that is publically available. Analyzing the public data first will give me a chance to build and refine a regression model that I can use with the daily data. It will also be useful to see if this is in fact a good method with which to study environmental health disparity in the Refinery Corridor.

Environmental Health Disparity in the Bay Area

The Bay Area, and in particular the city of Richmond, have been noted in the literature as a place where environmental racism and environmental health disparity exist. The city's residents of color disproportionately live near the refineries and chemical plants, which is noted in early works on environmental racism by pioneers of the idea, such as Robert Bullard (Bullard 1993a,b). Since the issue has been brought to national attention by Environmental Justice groups like West County Toxics Coalition, progress has been made to try to bring justice, but it has been limited. The situation where people of color are disproportionately exposed to toxic, industrial pollution still exists. A recent study showed 93% of respondents in Richmond were concerned about the link between pollution and health, and 81% were concerned about a specific polluter, mainly the Chevron Refinery (Brody et al. 2012). Despite the concerns of the community, there have been very few studies done on the health impacts that the citizens of Richmond and the rest of the Refinery Corridor are subject to. I will now present some of the research which has been done on this area, which provides basis for the hypothesis that refinery and other industrial pollution will increase the risk of negative health outcomes.

While the literature on the health impacts due to pollution in the Bay Area Refinery Corridor is sparse, most of it supports the fact that these large polluters significantly impact health. The study that looks at those that are the closest to refinery pollution is an epidemiological study on the workers at the Richmond refinery, and another refinery in Los Angeles, published by Satin et al. 2002. The study calculated standardized mortality ratios (SMR's), which compare how likely a subject (in this case, a refinery worker) is to die from a given health condition compared to the average person. The study found that workers at the Richmond refinery were more likely to die from cancers of the esophagus, larynx, rectum, and kidneys, skin, nervous system, and lymphatic tissues. There was a large and significant increase in the mortality rate for lymphosarcoma among those who had worked at the refinery for over 30 years. The rate of death from heart disease and stroke was lower than that in the general population, however this is likely due to the healthy worker effect. The healthy worker effect is the fact that those who work are likely to be healthier than the average person from the general population. This means that if you calculate SMR's of workers, you are likely to see that workers are less likely to die of most conditions than the average population. This does not indicate a protective effect of the work, but rather that work is a confounding factor.

Another study by Kaldor et al. 1984 examines pollution and health in the Refinery Corridor. They stratified cancer incidence rates and other cause specific mortality rates according to pollution exposure. They used the concentration of refineries and chemical plants to estimate pollution exposure, and stratified the areas in Contra Costa county into 4 groups, from highest concentration of polluters to the lowest. They found many similar things to the previous study, but their results were unaffected by the healthy worker effect so they are more applicable to the group we plan to study. In men, they found statistically significant increases in cancers of the mouth and throat, stomach, lungs, kidney, prostate, and overall cancers of all sites. In females, they found statistically significant increases in mouth and throat, esophagus and bladder cancers. They also found a statistically significant increase in cardiovascular mortality for both males and females as you move closer to the major polluters. This supports my expected findings, and also provides a good example for how I can perform my stratification, by concentration of major polluters in each zip code.

I found one more study which is relevant to ours, which examined the relationship between air pollution and health outcomes in 9 California counties, Contra Costa (the location of most of the Refinery Corridor) being one of them, published by Ostro et al. 2006. While their objective was not to examine any environmental justice related issues, I analyzed their results through that lense. Their results were not very informative for my purposes, mainly because they only used one monitor for the whole county. This means that those close to the refineries and those far away were analyzed using the same pollution data. Contra Costa did not have the highest association between air pollution and adverse health outcomes out of the counties they studied. However, many of the other counties examined were far more urban, so there are confounding factors. More importantly, this study's methods are a good guide for me to follow, since I plan to use many of the same techniques. These include calculating the relationship between pollution and health outcomes in many areas independently and also pooling them together. These are the only three studies I was able to find that directly examine the link between pollution and health in Contra Costa county. Together, they provide some evidence of an environmental health disparity, but more research is needed to provide a strong basis for this claim.

Hypothesized Mechanisms

The literature on pollution related health impacts in the Bay Area is very limited, despite the concerns of the residents there and the national attention this example of environmental racism has attracted. Despite this there is a large body of research into pollution based health impacts in general, and racial health disparity. My hypothesis that air pollution induced cardiovascular and respiratory health outcomes will differ inside and outside the refinery corridor is based on this research. I will now outline the research which I have based my hypothesis off of. The two main mechanisms I am expecting to cause this environmental health disparity are the speciation of the air pollution in these areas, and allostatic load.

PM 2.5 is defined as all the particulate matter in the air under 2.5 microns in diameter, however there are many different types of this particulate matter, and some of it is more dangerous to health than others. These particles can be made of a number of different chemicals, some less harmful, such as black carbon. They can also be composed of much more dangerous materials such as heavy metals or volatile organic compounds (VOC's). It is this distinction which gives reason to believe that a given amount of pollution could be more dangerous in the Refinery Corridor than in the surrounding areas.

The speciation of air pollution is a major factor in the relationship between ambient air pollution and the rate of health outcomes it triggers. PM 2.5 high in certain species, such as heavy metals and secondary species (those formed by reactions in the atmosphere, such as photochemical reactions) has been shown to trigger higher rates of hospital admissions than other species (Bell et al. 2009; Rich et al. 2013; Zanobetti et al. 2009). Industrial activity is known to be a principal source of heavy metal in the air (Chang 2003). Oil refining is also a major source of VOC's in air pollution, which can be involved in the creation of secondary species (Cetin et al. 2003). Speciation of the pollution coming from the Refinery Corridor may be enriched in particularly harmful species, which would explain how it causes more health impacts.

In addition to the fact that manufacturing and refining are known to produce particularly toxic pollution, research has been done on the specific makeup of pollution in the Refinery Corridor. The best study to do this is the Northern California Household Exposure Study (Brody et al. 2009). They examined indoor and outdoor air in Richmond, a Refinery Corridor community, and Bolinas, a nearby community which is far more rural. They found 33% more compounds in Richmond, along with higher concentrations of each compound. They found very high concentrations of vanadium and nickel in Richmond, some of the highest levels in the state. Vanadium and nickel have been shown to be some of the most dangerous PM 2.5 components as we previously stated, which gives reason to believe the pollution in Richmond is more toxic than in surrounding areas. Another very similar study compared the levels of endocrine disrupting compounds in Richmond and Bolinas homes, and found 40 in Richmond homes and only 10 in Bolinas (Rudel et al. 2010). While not directly related to short term pollution induced increases in hospital admissions, this supports the idea that the large variety of pollutants may be contributing to the increased toxicity of pollution in this area. This small body of research on pollution in Richmond suggests that the composition of air pollution may be more toxic and thus trigger more pollution related adverse health outcomes than in surrounding communities.

The other mechanism I believe may be at work here is the allostatic load of stress model. Allostatic load is defined as the cumulative "wear and tear" on the body due to a lifetime of adapting to stressful situations. It has been theorized that this stress builds up over time, as some stress responses overact to the stimulus, or perhaps fail to turn off (McEwen 1998). Markers of allostatic load such as blood pressure, hip to waist ratio and cortisol levels have been associated with higher mortality risks and declines in functioning (Seeman et al. 2001). If someone has a higher allostatic load, it is possible then that it will be easier for an additional insult, such as a day with particularly high air pollution, to set off an acute event such as a heart attack or stroke. This may be another contributing factor to why pollution in the Refinery Corridor is more likely to trigger one of these health outcomes.

Allostatic load has also been suggested as a possible mechanism of action for racial and socioeconomic health disparity. While the research to support this is limited, the allostatic load model offers a mechanism in which factors such as the stress which poverty causes, or the mental impact of facing discrimination can impact health. Allostatic load has been shown to be higher in those of lower socioeconomic status (Seeman et al. 2001). High allostatic load has also been associated with neighborhood poverty, independent of individual poverty status (Schulz et al. 2011). This research gives reason to believe that allostatic load may be elevated in the Refinery Corridor.

Conclusion

The environmental health disparity in the Refinery Corridor has not been well studied, despite being a common example of environmental racism. By studying the association between air pollution and acute cardiovascular health outcomes in this community, I hope to uncover one of the ways in which this pollution impacts the health of residents. Proof of this negative effect can be used by environmental justice groups to help bring about policy change to end this environmental health disparity. I also hope our adaptation of a traditional study design will encourage others to attempt to directly study the effect of pollution on communities, as well as environmental health disparities.

METHODS

Data Collection Methods

My mentor and I gathered the data on pollution and health outcomes for this study from two public agencies, the EPA and the California Department of Public Health. The data on pollution was downloaded from the EPA's Air Map website, and included comprehensive data from all PM 2.5 and SO2 monitors in Contra Costa and Solano counties for the years 2008-2015. This time period was selected based on the availability of daily air monitor data. The data was originally generated by the EPA using their monitoring equipment according to government protocol, which are outlined in 40 CFR Part 50 (Code of Federal Regulations). The pollution measure I am using in my analysis is the 24-hour block average. To calculate the 24-hour block average, the monitor calculates an hourly average every hour, and at the end of the day all of those are averaged together. In addition, a "Five-Minute Maximum" reading for SO2 was provided by the EPA. This is the highest five-minute block average that was recorded on each day, and it is the value I used for SO2 in the analysis. Some days had negative pollution readings those values were all changed to zero. Quarterly pollution averages were from these daily readings in order to match with the quarterly pollution averages. Information from 4 PM2.5 monitors and 7 SO2 monitors was used. Monitor Locations are shown in Figure 1.

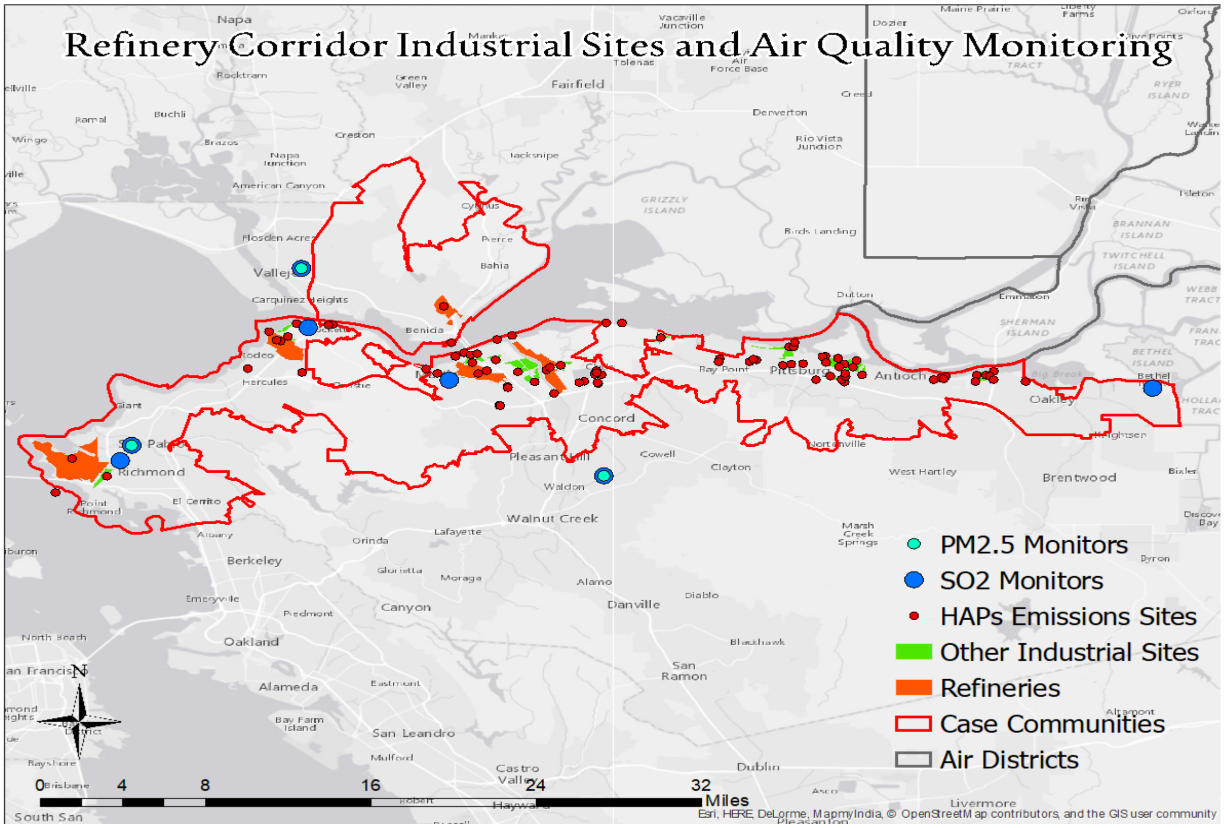

Figure 1. Monitor and Refinery Locations and Refinery Corridor Definition. Map showing the locations of PM 2.5 and SO2 monitors, refinery locations, and the definition of the Refinery Corridor I used.

Data on health outcomes was taken from the Office of Statewide Health Planning and Development's public ED use reports. The reports were accessed at http://www.oshpd.ca.gov/HID/Facility-Summary-Reports.html. They are Emergency Department Facility Summary Reports, which summarized ED admissions for 4 periods each year (quarterly reports). Each Emergency Department releases one of these reports, and all the hospitals in Contra Costa and Solano counties were included. Admissions counts from the Circulatory and Respiratory categories were collected, and rates were calculated using total ED admissions for each 3-month period. The formula to calculate rates for respiratory and circulatory admissions was:

Admission Rate = [Admission Count/Total ED Admissions]*10,000

Pollution and Demographic Analysis

In order to determine which zip codes were considered inside the Refinery Corridor and which were considered outside, I determined which census tracts were within a one mile buffer zone around each refinery. Any zip code which contained one of these census tracts was considered affected. Additionally, zip codes which were in the top 20 percentiles of toxic emissions according to Cal EnviroScreen were included. The areas that were considered inside the Refinery Corridor are shown in Figure 1. I then used census information to determine the demographic makeup of the two different areas. This analysis allowed insight into environmental injustice in the Refinery Corridor.

Analysis of Pollutant/Hospital Admissions Relationship

I examined the relationship between air pollution and hospital admissions for cardiovascular and respirator complaints by combining the two previously described data sets. Specifically, I used data from the closest air monitor to each hospital to determine the level of air pollution during each quarterly period. I examined the relationship between both pollutants and both classes of admission for a total of six pairs (PM 2.5, SO2, SO2 5-Minute Maximum, Respiratory rates, Circulatory rates). This analysis assumes that every patient lived closest to the hospital they were admitted to.

For each pollutant-outcome pair, I conducted a linear regression using the statistics program “R”. I also conducted a multiple linear regression and a mixed regression. I used log transformed data for all of these analyses because it was more normally distributed. To log transform the data, the formula was:

Log Transformed Value = Ln(Pollution level + 1)

The regression analysis allowed me to determine if higher levels of air pollution are predictive of higher rates of ED admissions. The independent variable being considered was air pollution, while the dependent variable was rate of admissions for CVE and asthma. For the multiple linear regression, both pollutants were used as independent variables, and season was considered as a factor. I used the 5-Minute Maximum value for SO2 instead of the 24 hour block average in the multiple regression models. Season was included because both the rate of respiratory and circulatory admissions, as well the levels of PM2.5 and SO2 vary seasonally. The formula for my regression model was:

Admission Rate ~ PM2.5 + SO2 5-Minute Maximum + Season(Factor)

I also constructed a mixed model. Since the hospital admissions are not necessarily independent (patients are either admitted to one hospital or another, and that decision is often determined by non-random factors like insurance or how crowded hospitals are), a mixed model is useful to help deal with these issues. In the mixed models, both air pollutants were considered fixed factors, and season and hospital were considered random factors. I included hospital in this analysis to determine if it should be included in the final model I use. I used the lme4 package in R to create the mixed model. The formula for the mixed model was:

Admission Rate ~ PM2.5 + SO2 5-Minute Maximum + (1|Season) + (1|Hospital)

where the (1|Season) syntax specifies that season is a random factor. The R-squared value and P-value were reported as results, where the P-value is the probability of obtaining the given correlation coefficient given that its true value is 0. The correlation was considered significant if the P value is below 5%. For the mixed models, variance is reported instead of R-squared. To calculate P-values for each pollutant in the mixed models, an ANOVA was used to compare the model including the pollutant to the null model which did not include the pollutant.

Spatial Variation of Relationship Between Air Pollution and Hospital Admissions

After analyzing the relationship over the whole study areas, I separated the Refinery Corridor from the surrounding area to see if the relationship differed between the two areas. Specifically, I was interested in examining whether the hospitals that are part of the Refinery Corridor have a stronger correlation between pollution and hospital admissions. I used the same models from the previous whole group analysis, but analyzed the hospitals inside the Refinery Corridor separately from those outside.

RESULTS

Demographic and Pollution Burden Analysis

The first analysis I performed was a demographic analysis of the Refinery Corridor, compared to the surrounding areas. I compared the demographics of the Refinery Corridor to the demographics of the rest of Contra Costa County. Data was derived from Census information, specifically the 2014 American Community Survey 5 year estimates. The demographics of the Refinery Corridor communities were different than the demographics of the surrounding communities. There were more people of color, less Caucasians and more people below the poverty line in the Refinery Corridor compared to the rest of Contra Costa county and the county as a whole. These demographics are summarized by county in Table 1.

I also analyzed the pollution burden inside and outside the Refinery Corridor (Table 2). PM2.5 was not particularly elevated in the Refinery Corridor. However, SO2 levels were higher inside the refinery corridor. This makes sense, because SO2 is released by refining, while PM2.5 is not. This supports my assumption that SO2 would be the pollutant that was elevated in the Refinery Corridor compared to the surrounding areas, not PM2.5.

Table 1. Demographics of study areas. This table shows the demographics composition of the study population.

Contra Costa County Demographics by Zip Code							
	Population	People of Color	Hispanic	White Only	\frican American	Asian	Poverty
Contra Costa	1,096,068	54.00%	24.90%	46.00%	10.60%	18.50%	10.30%
Refinery Corridor	611,655	67.95%	33.84%	32.16%	15.55%	18.47%	14.93%
Surrounding Areas	484,413	40.23%	14.17%	59.77%	5.95%	19.98%	3.85%

Table 2. Pollution Distribution. This table shows the levels of PM2.5 and SO2 at the monitors used in the study

Refinery Corridor			Surrounding Areas		
Monitor	PM 2.5	SO2	Monitor	PM 2.5	SO2
Richmond (Rumrill Blvd)	10.48	1.44	Concord	7.97	0.52
Richmond (7th St)	NA	0.73	Vallejo	9.79	0.69
Martinez	NA	1.18	Bethell Island	NA	0.89
Crockett	NA	1.29	Napa	11.47	NA

Pollution and health outcome regressions and correlations

Ambient air pollution levels were significantly correlated with health outcomes over the whole study area. When I aggregated the data from the whole study area and performed regressions on health outcome rates and pollution levels, I found that both PM 2.5 and SO2 were significantly and positively correlated to respiratory admissions, but negatively related to circulatory admissions. Table 3 shows the results of the linear regressions for each pollutant- outcome pair. The results of the multiple linear regression model were similar, with PM2.5 and SO2 positively and significantly associated with respiratory admissions, and negatively associated with circulatory admissions (Table 4).

Table 3. Linear regression results. The linear regression coefficients, R-squared values and P-values are shown for each pollutant-outcome pair.

All Areas- Log Transformed Linear Regression			
Respiratory Admissions			
	Coefficient	R-Squared	P Value
PM2.5	0.57	0.12	7.30E-11
SO2	0.29	0.03	0.00067
SO2 5-Minute Ma	0.23	0.02	0.016
Circulatory Admissions			
PM2.5	0.015	1.00E-04	0.847
SO2	-0.35	0.06	1.68E-06
SO2 5-Minute Ma	-0.33	0.078	8.86E-06

Table 4. Multiple Linear Regression Model. The regression coefficients, P values for each pollutant, and the P value and R-squared values for the overall model are shown.

All Areas- Log Transformed Multiple Linear Regression			
	Respiratory Admissions		
	Coefficient	R-Squared	P Value
PM2.5	0.35	NA	0.0004
SO2 5-Minute Maximu	0.17	NA	0.035
Overall Model	NA	0.53	2.2E-16
	Circulatory Admissions		
PM2.5	-7.00E-02	NA	0.49
SO2 5-Minute Maximu	-0.42	NA	4.73E-06
Overall Model	NA	0.08	2.20E-16

Lastly, the results of the mixed model are shown in Table 5. When both hospital and season are included as random factors, the relationship between respiratory admissions and both pollutants are not significant. The variance for season and hospital are both higher than the residual variance, showing that season and site explain most of the variance in this model. For circulatory admissions, PM 2.5 is not significantly related, and SO2 shows a significant negative correlation. The variance for season is much lower for circulatory admissions, while the variance for hospital is much higher. The residual variance is also much lower, meaning that the model for circulatory admissions explains much more of the variability in admissions than the model for respiratory admissions.

Table 5. Mixed Regression Model. The variance and standard deviation are shown for each random factor, and the coefficients and P value are shown for each fixed factor.

Full Mixed Model (lme4)					
Respiratory Admissions			Circulatory Admissions		
Random Effects			Random Effects		
	Variance	SD		Variance	SD
Hospital	0.054	0.23	Hospital	0.083	0.29
Season	0.093	0.31	Season	0.00071	0.027
Residual	0.021	0.14	Residual	0.009	0.095
Fixed Effects			Fixed Effects		
	Coefficient	P-Value		Coefficient	P-Value
PM2.5	0.08	0.23	PM2.5	0.019	0.64
SO2 5 Minute-Maximum	0.1	0.14	SO2 5 Minut	-0.102	0.03

Refinery Corridor vs Surrounding Areas Analysis

Analyzing the Refinery Corridor and surrounding areas separately with a multiple linear regression model revealed differences in the air pollution-health outcome relationship over the study area. PM 2.5 was not significantly related to respiratory or circulatory admissions in the Refinery Corridor, while SO2 was significantly negatively correlated. Outside the Refinery Corridor, PM2.5 and SO2 were positively and significantly related to respiratory admissions, contrary to my hypothesis. PM 2.5 was not significantly related to circulatory admissions, while SO2 was significantly negatively related. These results are summarized in Table 6.

Table 6. Refinery Corridor and surrounding areas comparison. This table contains the coefficients and P-values for each pollutant, as well as the R-squared values and P-values for the overall models. The Refinery Corridor and surrounding areas were analyzed separately.

Refinery Corridor vs Control-Multiple Linear Regression						
	Refinery Corridor			Control Areas		
	Respiratory Admissions					
	Coefficient	R-Squared	P Value	Coefficient	R-Squared	P Value
PM2.5	0.17	NA	0.15	0.57	NA	8.60E-09
SO2 5-Minute Maximum	-0.27	NA	0.003	0.5	NA	0.0003
Overall Model	NA	0.58	2.2E-16	NA	0.76	2.20E-16
	Circulatory Admissions					
PM2.5	0.063	NA	65	-0.033	NA	0.003
SO2 5-Minute Maximum	-0.023	NA	0.83	-0.47	NA	0.002
Overall Model	NA	-0.03	2.2E-16	NA	0.2	2.2E-16

DISCUSSION

This study examined whether the relationship between air pollution and health outcomes differs between the Refinery Corridor and the surrounding communities, to see if those in the Refinery Corridor are more affected by pollution induced health issues. To do this, I compared the rate of Emergency Department admissions for cardiovascular and respiratory events with air pollution data and analyzed it using a linear regression model. I found that the relationship did vary between hospitals located inside the Refinery Corridor and those outside, but contrary to my hypothesis. These regression results mean that PM2.5 and SO2 as measured by the EPA are not better predictors of ED visits in the Refinery Corridor compared to the surrounding communities. The higher R-squared values outside the Refinery Corridor indicate that there are more factors at play in determining the rate of respiratory and circulatory ED admissions in the Refinery Corridor than the surrounding areas.

Spatial Distribution of Demographics and Pollution Burden

Demographic analysis of our study area support the assertion that there is a racial and economic environmental disparity in this community. The communities which we included in our definition of the Refinery Corridor based on proximity to polluters had different demographics from the surrounding communities. The proportion of people of color was 68% in the Refinery

Corridor compared to 40% in the rest of Contra Costa County. 32% of people identified as non-Hispanic whites in the Refinery Corridor, while 60% identified that way in the rest of the county. The communities in the Refinery Corridor were also on average lower income than the surrounding communities. 15% of those living in the Refinery Corridor were below the poverty line, while only 4% of those in the rest of the county were below the poverty line. The demographic evidence supports the assertion that people of color and low income are more likely to be exposed to environmental hazards in Contra Costa County.

These results support a large body of literature which has identified the Refinery Corridor, and specifically the town of Richmond, as a place where people of color and those of low income are unjustly exposed to high levels of industrial and refinery pollution. Robert Bullard, one of the pioneers of the theory of environmental racism, has written extensively about environmental racism in Richmond (Bullard 1993). He has identified the fact that in the city of Richmond, the communities nearest to the Chevron Refinery are overwhelmingly made up of people of color. Other, more contemporary studies such as Brody et al. 2012 have also found that this is an area where marginalized populations are exposed to the most environmental hazards. My demographic findings show that the areas which are more heavily impacted by refinery pollution are indeed the areas with the highest concentrations of people of color and people of low income, vindicating claims by community advocated in these areas.

Ambient Air Pollution and Health Outcomes

The pooled analysis of all hospitals analyzed were consistent with previous studies which have shown that levels of PM 2.5 and SO2 are associated with increased acute respiratory health events, but inconsistent with studies on cardiovascular health events. The results of the linear regression relating the rate of ED visits with the average pollution indicated that PM2.5 and SO2 were only significantly positively related to respiratory ED visits, and not circulatory ED visits. The results of the multiple linear regression and mixed models showed the same results. The results of the mixed models identified the fact that using hospital as a random variable impacted the circulatory model much more than the respiratory model. This means that which hospital you sample impacts the relationship between pollutants and ED visits much more for circulatory visits than respiratory visits. Also, season had much less variance associated with it in the circulatory mixed model than the respiratory mixed model, indicating that season is much less of a factor in circulatory admissions. Lastly, the residual variance was much lower in the circulatory mixed model, meaning that season, hospital and the two pollutants explain much more of the variability in circulatory ED admissions than respiratory.

This analysis is an important contribution to the literature on the relationship between air pollution and health. This is one of the first studies I am aware of in America which specifically examines an environmental justice concern area. The relationship we found between cardiovascular ED visits and PM 2.5 was not as strong as the one found by Ostro et al. 2006 when they examined the relationship in Contra Costa county. This may be because they had access to daily data, while I did not for this study. Compared to the result presented in Dockery and Pope 1994, we found a stronger relationship between PM and respiratory visits. This supports my hypothesis that the speciation of PM in this area is more toxic due to the many industrial polluters. I saw evidence of a relationship between SO2 and acute health outcomes, which has not been as consistently seen as the one between PM 2.5. Poloniecki et al. 1997 found no significant relationship between heart attacks and SO2, while Katsouyani et al. 1997 did find a significant relationship. I found a significant relationship between SO2 and respiratory outcomes, but not circulatory outcomes. This may indicate that there are more factors influencing this relationship than simply the concentration of SO2, such as seasonal variations and the source of the pollution.

Differences Between Refinery Corridor and Surrounding Communities

The relationship between air pollution and adverse health outcomes was less strong in the Refinery Corridor, which is inconsistent with the hypothesis that air pollution and ED visits are more strongly associated in the Refinery Corridor than the surrounding communities. R-squared values for the linear regressions and multiple linear regressions were lower inside the Refinery Corridor for all pollutants and the multiple linear regression model. However, the fact that the R-squared values are higher outside the Refinery Corridor means that the two air pollutants and season explain much more of the variance in ED admissions outside the Refinery Corridor. This means that there may be more pollutants and factors that are influencing the relationship inside the Refinery Corridor that I have not accounted for. These could include other refinery pollutants like VOC's that were not measured on a daily basis by the EPA. Adding these to the model may change the results. Also, the demographic analysis revealed that the populations outside the Refinery Corridor were older, which means they are more at risk for these types of health issues. Using data that includes age will allow me to control for this.

The results I found are contradictory to the findings of Kaldor et al. 1984, who did one of the first studies looking at the relationship between pollution and health in this area. They found that cancer rates are higher in areas of Contra Costa county located closer to the refineries and chemical manufacturers. I found the opposite when I examined acute health effects and their relationship to pollution. A study by Brody et al. found higher pollutions levels and a wider variety of dangerous pollutants in homes in Richmond compared to rural Bolinas, which offers a possible mechanism for how this difference may be arising. These results show that the difference in the composition of pollution cannot be measured just by using PM 2.5 and SO2. My results overall do not provide evidence of environmental health disparity. More detailed health and pollution data will need to be analyzed to better understand how refinery pollution impacts these populations.

Limitations and Future Directions

The biggest limitation we faced in carrying out this research was the lack of daily air pollution data. The public data sets are not detailed enough to do the do a thorough study on the daily association between air pollution and cardiovascular and respiratory hospitalizations. I was also limited by the availability of air monitor data. There were 10 monitors for the two main pollutants of interest in the study area, which meant that many communities we were trying to compare had to be assessed using the same monitor. The EPA was the only publically available source of pollution data I could find, so this problem can be addressed by more independent monitoring. The other main limitation faced was the vast number of confounders that exist. Air pollution is far from the most influential factor affecting the rate of heart attacks at any given time. Temperature, time of year, age distributions, and many other factors all affect this relationship, and can overpower the effect of air pollution.

One of the main goals in conducting this study was to spur interest in studying environmental health disparities. I hope that my study can serve as a framework moving forward for studying how pollution affects communities who are exposed to refinery and industrial pollution, which is particularly dangerous. Adapting existing study designs is much easier than imagining new ones, and allows investigators to use tested methods known to work instead of brand new ones which may not be effective. Since this was an exploratory study, the results will help guide future research by identifying which directions hold promise. We were excited by our results regarding (list), and hope that others will follow up and further investigate how pollution impacts communities differently, especially those who are the most exposed to toxic, industrial pollution.

Broader Implications

The overarching goal of this study was to provide empirical evidence to support the claims of environmental justice groups and other community advocates who feel that they bear a disproportionate burden of the health impacts of pollution. The burden of proof is often on these communities, and they are the ones who have to prove that their health is jeopardized by these large polluters. The lower association between PM 2.5 and SO2 in the Refinery Corridor compared to the surrounding communities I found shows that this relationship is more complex than just the relationship between pollution concentrations and incidence rates. I hope that more research using a more detailed version framework as well as other study designs can be used to build a large body of quantitative literature on environmental health disparity. This will allow these communities to drive policymaking and political action regarding pollution and environmental hazards which is an important and necessary step towards a safe environment for all.

ACKNOWLEDGEMENTS

Kurt Spreyer, Patina Mendez, and Abigail Cochran gave me the instruction and guidance that allowed me to create this thesis. I am so grateful for the time Abby took to read my work and give me constant feedback that allowed me to become a better writer. My mentor Kyle Ferrar also provided me with guidance and worked with me on many aspects of this project. I am so thankful that he reviewed my work, and spent countless hours working with me on this project. Howard Kipen and Kathleen Black introduced me to the field of environment health, and helped me with the initial conception and visualization of this project. My peer review partner Kristy Hung reviewed every piece of writing I did and always offered constructive criticism.

REFERENCES

Bell, M. L., K. Ebisu, R. D. Peng, J. M. Samet, and F. Dominici. 2009. Hospital Admissions and Chemical Composition of Fine Particle Air Pollution. American Journal of Respiratory and Critical Care Medicine 179:1115–1120.

Brody, J. G., R. Morello-Frosch, A. Zota, P. Brown, C. Pérez, and R. A. Rudel. 2009. Linking Exposure Assessment Science With Policy Objectives for Environmental Justice and Breast Cancer Advocacy: The Northern California Household Exposure Study. American Journal of Public Health 99:S600–S609.

Brook, R. D., B. Franklin, W. Cascio, Y. Hong, G. Howard, M. Lipsett, R. Luepker, M. Mittleman, J. Samet, S. C. Smith, and I. Tager. 2004. Air Pollution and Cardiovascular Disease. Circulation 109:2655–2671.

Brook, R. D., S. Rajagopalan, C. A. Pope, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith, L. Whitsel, and J. D. Kaufman. 2010. Particulate Matter Air Pollution and Cardiovascular Disease. Circulation 121:2331–2378.

Bullard, R. D. 1993a. Race and Environmental Justice in the United States Symposium: Earth Rights and Responsibilities: Human Rights and Environmental Protection. Yale Journal of International Law 18:319–336.

Bullard, R. D. 1993b. Confronting Environmental Racism: Voices from the Grassroots. South End Press.

Cohen, A., A. Lopez, N. Malloy, and R. Morello-Frosch. 2012. Our Environment, Our Health: A Community-Based Participatory Environmental Health Survey in Richmond, California. Health Education & Behavior 39:198–209.

Cetin, E., M. Odabasi, and R. Seyfioglu. 2003. Ambient volatile organic compound (VOC) concentrations around a petrochemical complex and a petroleum refinery. Science of The Total Environment 312:103–112.

Dockery, D. W., and C. A. Pope. 1994. Acute Respiratory Effects of Particulate Air Pollution. Annual Review of Public Health 15:107–132.

Dominici, F., R. D. Peng, M. L. Bell, L. Pham, A. McDermott, S. L. Zeger, and J. M. Samet. 2006. Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases. JAMA 295:1127–1134.

Kaldor, J., J. A. Harris, E. Glazer, S. Glaser, R. Neutra, R. Mayberry, V. Nelson, L. Robinson, and D. Reed. 1984. Statistical association between cancer incidence and major-cause mortality, and estimated residential exposure to air emissions from petroleum and chemical plants. Environmental Health Perspectives 54:319–332.

Katsouyanni, K., G. Touloumi, C. Spix, J. Schwartz, F. Balducci, S. Medina, G. Rossi, B. Wojtyniak, J. Sunyer, L. Bacharova, J. P. Schouten, A. Ponka, and H. R. Anderson. 1997. Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. BMJ 314:1658.

McEwen, B. S. 1998. Stress, Adaptation, and Disease: Allostasis and Allostatic Load. Annals of the New York Academy of Sciences 840:33–44.

Ostro, B., R. Broadwin, S. Green, W.-Y. Feng, and M. Lipsett. 2006. Fine Particulate Air Pollution and Mortality in Nine California Counties: Results from CALFINE. Environmental Health Perspectives 114:29–33.

Poloniecki, J. D., R. W. Atkinson, A. P. de Leon, and H. R. Anderson. 1997. Daily Time Series for Cardiovascular Hospital Admissions and Previous Day's Air Pollution in London, UK. Occupational and Environmental Medicine 54:535–540.

Rich, D. Q., H. Özkaynak, J. Crooks, L. Baxter, J. Burke, P. Ohman-Strickland, K. Thevenet-Morrison, H. M. Kipen, J. Zhang, J. B. Kostis, M. Lunden, N. Hodas, and B. J. Turpin. 2013. The triggering of myocardial infarction by fine particles is enhanced when particles are enriched in secondary species. Environmental science & technology 47.

Rudel, R. A., R. E. Dodson, L. J. Perovich, R. Morello-Frosch, D. E. Camann, M. M. Zuniga, A. Y. Yau, A. C. Just, and J. G. Brody. 2010. Semivolatile Endocrine-Disrupting Compounds in Paired Indoor and Outdoor Air in Two Northern California Communities. Environmental Science & Technology 44:6583–6590.

Satin, K. P., W. J. Bailey, K. L. Newton, A. Y. Ross, and O. Wong. 2002. Updated epidemiological study of workers at two California petroleum refineries, 1950–95. Occupational and Environmental Medicine 59:248–256.

Schulz, A. J., G. Mentz, L. Lachance, J. Johnson, C. Gaines, and B. A. Israel. 2012. Associations Between Socioeconomic Status and Allostatic Load: Effects of Neighborhood Poverty and Tests of Mediating Pathways. American Journal of Public Health 102:1706–1714.

Schwartz, J., and R. Morris. 1995. Air Pollution and Hospital Admissions for Cardiovascular Disease in Detroit, Michigan. American Journal of Epidemiology 142:23–35.

Seeman, T. E., B. S. McEwen, J. W. Rowe, and B. H. Singer. 2001. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. Proceedings of the National Academy of Sciences 98:4770–4775.

Zanobetti, A., M. Franklin, P. Koutrakis, and J. Schwartz. 2009. Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environmental Health 8:58.