Air Pollution Sources and Childhood Asthma Attacks in Cataño, Puerto Rico

Nilsa I. Loyo-Berríos¹, Rafael Irizarry², Joseph G. Hennessey³,4, Xuguang Grant Tao⁵,⁶,⁷, and Genevieve Matanoski⁶,⁷

¹ Office of Surveillance and Biometrics, Center for Devices and Radiological Health, Food and Drug Administration, US Department of Health and Human Services, Rockville, MD.
² Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD.
³ Center for Imaging Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD.
⁴ Department of Psychiatry, School of Medicine, Johns Hopkins University, Baltimore, MD.
⁵ Division of Environmental and Occupational Health, School of Medicine, Johns Hopkins University, Baltimore, MD.
⁶ Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD.
⁷ District of Columbia Cancer Registry, District of Columbia Health Department, Washington, DC.

Received for publication May 9, 2006; accepted for publication October 9, 2006.

Asthma prevalence in the Cataño Air Basin of Puerto Rico is 27% for children aged 13–14 years and 45% for children aged 5–6 years. There is concern that these rates are related to air pollution. The authors conducted a nested case-control study to evaluate whether proximity to air pollution point sources was associated with increased risk of asthma attacks. For 1997–2001, 1,382 asthma-related medical visits (International Classification of Diseases, Ninth Revision, codes 493 and 493.9) in children under 17 were identified through health insurance claims. Controls were children with no asthma attacks who were randomly selected from enrollees in two health insurance companies by incidence density sampling (1:5) and matched to cases on gender, age, insurance company, and event date. The distance from a point source to the subject’s residence area represented a surrogate exposure measurement. Odds ratios for a 1-km decrease in distance were obtained by conditional logistic regression. Risk of asthma attack was associated with residing near a grain mill (odds ratio (OR) = 1.35), petroleum refinery (OR = 1.44), asphalt plant (OR = 1.23), or power plant (OR = 1.28) (all p’s < 0.05). Residence near major air emissions sources (>100 tons/year) increased asthma attack risk by 108% (p < 0.05). These results showed that proximity to some air pollution sources is associated with increased risks of asthma attacks.

Abbreviations: CI, confidence interval; OR, odds ratio.

Asthma is a common disease that affects people of all ages. It is a chronic inflammatory disorder of the lungs characterized by episodic and reversible symptoms of airflow obstruction and nonspecific airway hyperactivity (1, 2). In the United States, the self- or proxy-reported 12-month prevalence of asthma increased 73.9 percent from 1980 to 1996, when approximately 14.6 million persons (54.6 per 1,000 population) reported that they had had an asthma attack during the previous 12 months (3). In 1997, the National Health Interview Survey was redesigned, and the new estimate for 12-month prevalence was 40.7 per 1,000 population. However, an insufficient number of years with the new measures have passed to determine whether the trends in asthma are increasing or decreasing (3).

Ambient air pollution has been related to an increased risk of asthma attacks. Temporal associations have been shown...
between increases in the number of asthma attacks (emergency room visits/hospital admissions) and increases in levels of ambient air pollution in different parts of the world (4–12). Other studies have shown spatial associations, with asthma attacks being more common in areas that are closer to industrial parks, air pollution point sources, and high-traffic-volume roads (13–17).

Ambient air pollution and its health effects have not been studied in detail in Puerto Rico. The Cataño Air Basin, located west of San Juan Bay, is known for its air pollution problems and high asthma prevalences among children aged 6–7 years (45 percent) and 13–14 years (27 percent) (18). The area has been identified by the federal and local governments as a nonattainment area with regard to Environmental Protection Agency air quality standards for total suspended particulates and particulate matter less than 10 microns in diameter. The Cataño Air Basin includes the Cataño municipality (30,071 total population) and portions of four nearby municipalities: San Juan (434,374 total population), Guaynabo (100,053 total population), Bayamón (224,044 total population), and Toa Baja (94,085 total population) (19). The combined total population living in the Cataño Air Basin has been estimated at approximately 300,000 (20). Within its boundaries, there are numerous industrial point sources of air pollution, including a rum distillery, electric power plants, petroleum refineries, a sewage incinerator, sewage treatment plants, grain mills, and cement plants. Additionally, mobile sources include motor vehicle emissions from two highways and emissions from activities in a port zone. These air pollution sources are mostly concentrated in the Cataño municipality and its surrounding areas, generating concern about the observed high asthma prevalence rates being associated with the ambient air pollution.

Therefore, we conducted a study to evaluate whether there was an association between proximity to air pollution point sources and risk of asthma attacks in the Cataño municipality. We hypothesized that proximity to the air pollution point sources increased the risk of asthma attacks.

MATERIALS AND METHODS

Data sources

Two health insurance companies that operate in Cataño (referred to as company A and company B) provided health and demographic data on the study participants. Data on emergency room visits, hospital admissions, and visits to physicians’ offices were included in the study. Each company provides services to different sectors of the population (figure 1); company A provides services to people who can afford its private health insurance, and company B contracts with the government of Puerto Rico to provide services to different sectors of the population. Physicians’ offices were included in the study. Each company provides services to different sectors of the population (referred to as company A and company B) provided health care data on the study cohort and control cohort. Data on weather conditions (unpublished data), and census block maps were obtained from the US Census Bureau (24).

Study area and population

Puerto Rico is located in the Caribbean, and its weather conditions do not vary greatly across seasons of the year. The average temperature is 27°C (−80°F), and humidity averages 77 percent (National Climatic Data Center, unpublished data). The Cataño municipality is located in northern Puerto Rico, west of San Juan Bay; it is a small municipality with an area of 5.16 km² (24).

The study population consisted of children younger than 17 years who were residents of the Cataño municipality from January 1997 through December 2001. By the year 2000, they represented 30 percent of Cataño’s total population. The member rosters of the two health insurance companies included 69 percent of the child population through the year 2000 (figure 1).

Study design

We conducted a nested case-control study to evaluate the study hypothesis. This study design can be used when a cohort is available; it combines the features and advantages of both cohort and case-control designs, and because the cases and controls are selected from the same source cohort, selection bias is minimized (25). The study period consisted of 5 years of follow-up, starting in January 1997 and continuing through December 2001, and the two health insurance companies provided data on the cohort of children to be followed during that period. This cohort was dynamic; new children could enroll or children could leave the cohort, either because they became older than the age criteria or because their parents changed their health insurance company.

All of the cases observed during the follow-up period were included in the nested case-control study, and incidence density sampling was used to select the controls. Use of this sampling technique allowed for the selection of...
controls who were at risk and available in the cohort at the time an asthma attack was observed.

**Definitions of cases, controls, and exposure**

**Cases.** An asthma attack was counted as one case. Multiple attacks in the same child were considered separate cases. An incident attack was defined as an attack that occurred on any day for which the previous day was free of events. Asthma attacks were identified through insurance claims that had an asthma diagnosis made by a physician on the basis of symptoms and were coded from the *International Classification of Diseases*, Ninth Revision, as 493.0 (extrinsic asthma) or 493.9 (asthma, unspecified). Because this study focused on childhood asthma, the code used to indicate late onset of asthma (493.1) was not included. A total of 1,382 cases were identified during the study period.

**Controls.** Three control groups were used in the analysis. The first control group (control group 1) was defined as children who were free of an asthma attack on the day a case was observed. This group included asthmatic children and children who could have had other respiratory conditions (e.g., any respiratory condition other than asthma, such as "cataarrh, catarrhal (inflammation)" or "croup") on the day of the case. The second control group (control group 2) was defined similarly to the first group, except that children with other respiratory conditions on the day of the case attack were not included. Finally, the third control group (control group 3) consisted of children who never had an asthma attack or other respiratory condition during the 5 years of follow-up. Controls were matched to the cases on gender, age, insurance company, and date of the asthma event. Incidence density sampling was used to select five controls per case. A total of 6,910 controls were included in each control group.

**Exposure.** The physical distance from the air pollution point sources to the center of the census block in which the child’s residence was located was used as a surrogate measure for average exposure to ambient air pollution. A global positioning system was used to obtain the latitudes and longitudes of the center of each census block. Then, distance was calculated with R software (26) using the Haversine Formula (equations 1–5), which takes into account the curvature of the earth (27).

\[
\text{Radius} = 6,378 \text{ (km)}. \\
d\text{lon} = \text{lon2} - \text{lon1}. \\
d\text{lat} = \text{lat2} - \text{lat1}. \\
a = \left[ \left( \sin \left( \frac{d\text{lat}}{2} \right) \right)^2 + \cos(\text{lat1}) \times \cos(\text{lat2}) \right] \times \left( \sin \left( \frac{d\text{lon}}{2} \right) \right)^2. \\
c = 2 \times a \tan \left( \sqrt{a}, \sqrt{1-a} \right). \\
\text{Distance} = \text{radius} \times c.
\]

\( d\text{lat}, d\text{lon}, \text{lat1}, \text{lat2}, \text{lon1}, \text{lon2} \) are the residence area coordinates, \( \text{lon2} \) and \( \text{lat2} \) are the point source coordinates, and \( a \) is the argument inverse tangent. Distance is expressed in the same units as the earth’s radius. This calculation produced the simple distance from each point source to the residence area of each child.

**Statistical procedures**

The geographic distribution of the point sources was graphed using census maps in MapInfo Professional, version 7 (24).

Conditional logistic regression models were used to assess the association between distance from the residence to point sources and risk of asthma attack. R software, version 1.6.2, was used to conduct the analyses (26). The risk of asthma attack was estimated for a 1-km decrease in the distance to air pollution point sources. Three measurements of distance were used.

First, we used data from the Envirofacts Data Warehouse (21). The Envirofacts Data Warehouse classifies air pollution point sources as being major (≥100 tons/year) or minor (<100 tons/year) sources of air emissions (based on the total air emissions per year). That classification was used to estimate the mean distance from the residence to major or minor air emissions sources for each child.

Second, we also classified point sources according to types of industries that have been related to respiratory problems (power plants, grain mills, refineries, distilleries, asphalt plants, and quarries). We conducted a subanalysis for this exposure measurement using distance categories.

Third, we developed a corrected distance measurement to take wind direction into account (equation 6) (28).

\[
\text{Corrected distance} = \sum_{i=1}^{n} \left[ \left\{ \cos \left( \frac{\beta_i}{2} \right) \right\}^2 \times D_i \right], \quad (6)
\]

where \( D_i \) is the distance between the source, and the residence and \( \beta_i \) is the angle between the vectors for the wind direction and the source-residence vector (figure 2). The cosine function is maximal (equal to 1) when the wind is blowing in the same direction as the source-residence vector, and it is minimal (equal to 0) when the average wind direction is opposite that of the source-residence vector. The angles \( \beta_i \) are divided by 2 to ensure that the function will range only from 0 to 1, and it is squared to increase the drop-off rate when the wind vector and the source-residence vector are opposed. The distance from the residence to the source, is multiplied by the result of the cosine function. The distance is then expressed in metric units. Figure 2 presents an example of how the formula works if there are two sources of the same industry type (i.e., two power plants). Corrected distance measurements were estimated for two asphalt plants, two power plants, four grain mills, and five petroleum refineries.
RESULTS

Exploratory analysis

Figure 3 presents the geographic distribution of the point sources, as well as their classification by air emission type. There were 67 sources identified in the area, from which 20 were located within the municipality and 43 in the surrounding areas. There were no latitude and longitude data for four industries. As reported in the Envirofacts database (21), 17 sources were classified as major air emissions sources and 44 were identified as minor air emissions sources. Information on emission type was not available for six industries; those are represented as minor sources on the map and were excluded from risk estimation. Most of the major air emissions sources were located southeast of town. Two power plants, seven grain mills, five petroleum refineries, two asphalt plants, one quarry, and one rum distillery were identified in the area.

Multivariate analysis

Odds ratios were estimated for a 1-km decrease in the mean distance to major or minor air pollution sources. Table 1 presents the results. When sources were evaluated independently as major or minor air emissions sources, a 1-km decrease in the distance to major sources was associated with a 69 percent increase in risk (odds ratio (OR) = 1.69, 95 percent confidence interval (CI): 1.50, 1.91), and minor sources were associated with a risk increase of 47 percent (OR = 1.47, 95 percent CI: 1.32, 1.63). However, after controlling for type of air emissions, minor sources were no longer important, whereas major sources presented a 108 percent increased risk (OR = 2.08, 95 percent CI: 1.61, 2.70) (table 1). The observed associations were similar with the comparisons for control groups 2 and 3.

Table 2 presents the results for mean distance by type of industry. With the exception of power plant 1 (OR = 0.76, 95 percent CI: 0.72, 0.81) and the rum distillery (OR = 0.78, 95 percent CI: 0.73, 0.84), all of the sources were associated with increased risks of asthma attack; odds ratios ranged from 1.28 to 1.44, and the 95 percent confidence intervals excluded 1. The associations were similar for the comparisons with control groups 2 and 3. No effect modification was observed by wind speed or direction, temperature, humidity, or air pollution (data not shown). Figure 4 presents the results for the subanalysis by distance category and type of...
Children living within 2,000 m of the power plants and the rum distillery did not have increased risk. Interestingly, the risk estimates for the two power plants differed; the plant located north of town (power plant 1) never posed increased risk, whereas for the plant located southeast of town (power plant 2), the risk increased more rapidly and was higher for children living within 2,001–3,000 m. Furthermore, for all of the sources located to the south and west of the city, increased risk was observed for children living within 2,001–3,000 m.

**FIGURE 3.** Distribution of the air pollution point sources in Cataño and nearby municipalities, Puerto Rico, 1997–2001. Names of municipalities are shown in boldface type.

### TABLE 1. Adjusted odds ratios* for asthma attack per 1-km decrease in the mean distance from a major or minor industrial point source of air pollution to the subject’s area of residence (n = 1,382 cases†), Cataño municipality, Puerto Rico, 1997–2001

<table>
<thead>
<tr>
<th>Model and type of point source</th>
<th>Comparison†</th>
<th>Cases vs. control group 1</th>
<th>Cases vs. control group 2</th>
<th>Cases vs. control group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR§ 95% CI§</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Independent effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major source</td>
<td>1.69</td>
<td>1.50, 1.91</td>
<td>1.69</td>
<td>1.50, 1.91</td>
</tr>
<tr>
<td>Minor source</td>
<td>1.47</td>
<td>1.32, 1.63</td>
<td>1.49</td>
<td>1.34, 1.65</td>
</tr>
<tr>
<td>Adjusted for air emissions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major source</td>
<td>2.08</td>
<td>1.61, 2.70</td>
<td>1.95</td>
<td>1.51, 2.53</td>
</tr>
<tr>
<td>Minor source</td>
<td>0.81</td>
<td>0.64, 1.03</td>
<td>0.86</td>
<td>0.68, 1.09</td>
</tr>
</tbody>
</table>

* Odds ratios were obtained by conditional logistic regression. Cases and controls were matched on gender, age, health insurance company, and event date.
† 1:5 case:control ratio.
‡ See “Materials and Methods” for definition of control groups.
§ OR, odds ratio; CI, confidence interval.
southeast, proximity was associated with increased risks of asthma attack, and the risks decreased with distance.

Table 3 presents the results for the corrected distance models. The association observed for the grain mills, the petroleum refineries, and the asphalt plants remained after wind correction (tables 2 and 3). However, the corrected distance model showed a significant 12 percent increased risk for the power plants (OR = 1.12, 95 percent CI: 1.03, 2007;165:927–935

FIGURE 4. Adjusted odds ratios for asthma attack in children younger than 17 years per 1-km decrease in the mean distance from air pollution point sources to area of residence, by type of industry and distance category, Catanó municipality, Puerto Rico, 1997–2001. The maximum distance for power plant 1 was 4,546 m (two participants grouped in the distance category 3,001–4,000 m); the minimum distance for the petroleum refineries was 2,058 m; and the maximum distance for the rum distillery was 3,938 m. Cases and controls were matched on age, gender, health insurance company, and event date.
DISCUSSION

The study hypothesis postulated that proximity to industrial point sources of air pollution could increase the risk of childhood asthma attacks.  Various measurements of distance were used as indicators of average exposure to ambient air pollution. Results appeared to be consistent: Proximity to some point sources was related to increased risk of asthma attacks. The highest risk estimates were observed for point sources located in the south and southeast.

The results of our study are supported by reports in the scientific literature suggesting that proximity to air pollution point sources is related to increased risk of respiratory health problems (13, 15–17). A study conducted in northeastern England showed that persons living closest to air pollution sources had higher risks of having had more than 12 asthma attacks in the previous year (17). In another study in West Virginia, investigators found a higher risk of respiratory symptoms among children attending schools in a valley where industrial sources were located (13).

In the present study, we not only evaluated distance as a risk factor but also identified some other aspects that can influence the spatial association between air pollution and asthma attack risk. These factors are the amount of emissions and the location of the point sources. The data showed that major air emissions sources strongly influenced the risk estimate, with a 108 percent increase in risk. These data suggest that if those 17 sources identified as major air polluters by weight of output were able to decrease their annual emissions, the risk of asthma attacks in the Cataño Air Basin could be minimized. Additionally, the location of the sources proved to be another important factor determining risk.  The sources that were located north of town did not present increased risks. This was not the case for sources located to the south and southeast, all of which posed increased risks. Furthermore, this analysis identified two sources of the same type (power plant 1 and power plant 2) that presented opposite risks. The only difference between these two sources was their location; the power plant located in the southeast presented a risk that increased by 28 percent for each 1-km decrease in distance, whereas the power plant located in the north did not present increased risk.

Wind direction could be the reason for the high risk observed in the southeast. In this area, the wind mostly blows from the northeast, but there are times when it blows from the east. It is possible that emissions from the sources located in the southeast are carried towards the population living in the southeastern part of town. One of the strengths of this study is that a measurement of distance was developed to account for the wind direction on the day of the event (an asthma attack) (figure 2). This model assumes that same-day exposure will trigger an attack. After wind correction, the results changed only for the power plants, which presented a 12 percent increase in the risk associated with a 1-km decrease in distance.

Other factors that may also influence the risk are weather conditions (temperature and humidity), parents’ occupations, and where children go to school. Weather conditions in Puerto Rico remain similar across all seasons of the year. Therefore, weather factors were considered to be relatively constant over the 5 years of study. In fact, analyses stratified by temperature and humidity did not show effect modification for the observed association between distance and increased risk of asthma attack. On the other hand, parents’ occupations may affect the risk estimate if their working environments involve exposure to contaminants that can trigger asthma attacks and if those contaminants are brought home on the parents’ clothing. This study did not include data on parents’ occupations; further research examining indoor air quality and related factors is warranted. Additionally, where children go to school is an important factor to consider, since children spend most of their time at school.

---

**TABLE 3. Adjusted odds ratios* for asthma attack per 1-km decrease in mean distance from air pollution point sources to the subject’s area of residence (n = 1,382 cases†), by type of industry, with correction for same-day wind direction, Cataño municipality, Puerto Rico, 1997–2001**

<table>
<thead>
<tr>
<th>Type of industry</th>
<th>Cases vs. control group 1</th>
<th>Cases vs. control group 2</th>
<th>Cases vs. control group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR§ 95% CI§</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Power plants</td>
<td>1.12 1.03, 1.22</td>
<td>1.10 1.01, 1.19</td>
<td>1.10 1.01, 1.20</td>
</tr>
<tr>
<td>Grain mills</td>
<td>1.38 1.28, 1.49</td>
<td>1.37 1.27, 1.48</td>
<td>1.39 1.29, 1.50</td>
</tr>
<tr>
<td>Petroleum refineries</td>
<td>1.33 1.24, 1.42</td>
<td>1.34 1.25, 1.43</td>
<td>1.34 1.26, 1.44</td>
</tr>
<tr>
<td>Asphalt plants</td>
<td>1.39 1.29, 1.50</td>
<td>1.39 1.29, 1.50</td>
<td>1.41 1.31, 1.52</td>
</tr>
</tbody>
</table>

* Odds ratios were obtained by conditional logistic regression using one type of industry per model. Cases and controls were matched on gender, age, health insurance company, and event date.
† 1:5 case:control ratio.
‡ See “Materials and Methods” for definition of control groups.
§ OR, odds ratio; CI, confidence interval.
This study considered residence area rather than school location, and information on school location was not available. However, the study area is a small municipality, the school system in Puerto Rico assigns children to schools by residence area, and the norm is for children to go to schools within their residence area. Very rarely will a student go to a school that is either outside of his/her residential area or outside the municipality. Therefore, in this study, we considered residence area an appropriate indirect measurement for average exposure to outdoor air pollution.

Another aspect to consider when interpreting these results is that health insurance claims data capture only those children who need to seek health services. In this study, it was possible that children who had an asthma attack but had medications at home and did not seek health services for their attack were not captured as cases. Data on medications were not available for this study, and missing cases or misclassification of cases as controls may have resulted in lower odds ratios. We addressed this issue by using three different control groups: Control groups 1 and 2 included asthmatic children with a history of attacks, and control group 3 excluded asthmatic children. Control groups 1 and 2 should have included children with a history of asthma attacks and medications at home. The risk estimates for these two groups will tend to have been lower if cases were missed or if case children were misclassified as controls because of self-medication. Note that self-medication may not fully treat severe attacks, so some children with medications at home may still have needed to seek health services and therefore would have been captured as cases. Note also that the observed odds ratios showed increased risk, and they could have been underestimated. Future research considering medication use is warranted.

Another limiting aspect of these data was the diagnosis of asthma for children younger than 1 year of age. Asthma diagnosis for this age group is less specific, which could lead to misclassification of controls as cases. In this study, only 3.9 percent of the cases were younger than 1 year (data not shown); therefore, we would expect the effect of such misclassification to have been minimal.

This study had several other limitations. Correlation between locations of point sources prevented the identification of a specific source as the sole cause of increased asthma attack risk. When distance to an independent source is used, the estimated risk carries the risk of nearby sources. In this study, we addressed this limitation by classifying point sources as major or minor air emission sources, irrespective of their location. After controlling for emission type, increased risk was associated with proximity to point sources with major air emissions. Most of these major polluters were located south and southeast of town.

Finally, there was a lack of data on smokestack height, temperature of emissions, and emissions rate, which could have been used to determine how the emissions were distributed in the area and affecting the community. Nevertheless, this study showed that when these data are not available, valid estimates of risk can be obtained. Not only can this method be used in communities with similar health and air pollution problems, it can also be used to examine other health outcomes and exposures.

In future research in Cataño, investigators should consider evaluating how emissions are distributed and how they affect the community, perhaps using the Gaussian Plume model (29). Special attention should be given to the two power plants. Evaluations of the point sources located south and southeast of town are recommended, especially those classified as major air emissions sources. Minimizing emissions from these sources could reduce the occurrence of asthma attacks in this population.

ACKNOWLEDGMENTS

This research was conducted to fulfill the requirements for a doctoral degree in epidemiology and was funded by the Department of Epidemiology at the Johns Hopkins Bloomberg School of Public Health (Baltimore, Maryland). Dr. Nilsa I. Loyo-Berrios currently works for the Food and Drug Administration and prepared this paper under the Professional Development Program.

The authors thank Dr. Luis Marini, Dr. Benjamin Santiago, Carlos Morel, and Dr. Alejandro Franco, from the health insurance companies that provided the data on asthma. They also acknowledge Dr. Erika Avila-Tang for her help with data abstraction.

Conflict of interest: none declared.

REFERENCES