Using air pollution based community clusters to explore air pollution health effects in children

Zhengmin Qian\textsuperscript{a,b,*}, Robert S. Chapman\textsuperscript{c}, Wei Hu\textsuperscript{d}, Fusheng Wei\textsuperscript{d}, Leo R. Korn\textsuperscript{e}, Junfeng (Jim) Zhang\textsuperscript{b}

\textsuperscript{a}Department of Health Evaluation Sciences, Pennsylvania State University College of Medicine, 600 Centerview Drive, P.O. Box 855, Hershey, PA 17033, USA
\textsuperscript{b}Environmental and Occupational Health Sciences Institute, University of Medicine and Dentistry of New Jersey and Rutgers University, 170 Frelinghuysen Road, Piscataway, NJ 08854, USA
\textsuperscript{c}National Center for Environmental Assessment, US Environmental Protection Agency, Mail Drop 8243-01, Research Triangle Park, NC 27711, USA
\textsuperscript{d}China National Environmental Monitoring Center, Beisihuandonglu, Chaoyang District, Beijing 100029, China
\textsuperscript{e}New Jersey Department of Environmental Protection, Division of Science, Research and Technology and the Division of Biometrics, UMDNJ-SPH, 401 East State Street, Trenton, NJ 08625-0409, USA

Received 20 July 2003; accepted 10 November 2003

Abstract

To study respiratory health effects of long-term exposure to ambient air pollutant mixture, we observed 7058 school children 5–16 years of age living in the four Chinese cities of Lanzhou, Chongqing, Wuhan, and Guangzhou. These children were enrolled from elementary schools located in eight districts, one urban district and one suburban district in each of the above cities. Ambient levels of PM\textsubscript{2.5}, PM\textsubscript{10-2.5}, total suspended particles (TSP), SO\textsubscript{2}, and NO\textsubscript{x} were measured in these districts from 1993 to 1996. Based on a cluster analysis of arithmetic mean concentrations of PM\textsubscript{2.5}, PM\textsubscript{10-2.5}, (TSP–PM\textsubscript{10}), SO\textsubscript{2}, and NO\textsubscript{x}, we classified these children into four ordinal categories of exposure to ambient air pollutant mixtures. We tested for exposure–response relationships using logistic regression models, controlling for relevant covariates. We observed monotonic, positive relationships of exposure to the pollutant mixture with prevalence rates of cough with phlegm and wheeze. Other outcomes were not associated with the exposure in a monotonic exposure–response pattern. Even so, odds ratios for cough, phlegm, bronchitis, and asthma in the higher exposure district clusters were all higher than in the lowest exposure district cluster. We found evidence that exposure to the pollutant mixtures had adverse effects on children living in the four Chinese cities.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Air pollution; Health effects; China

1. Introduction

High ambient concentrations of air pollutants have been commonly present in many Chinese cities (Waldman et al., 1991; He et al., 1993; Zhang et al., 1999; Wei et al., 1999). In the four Chinese cities of Chongqing, Guangzhou, Lanzhou, and Wuhan, for example, we observed that pollution ranges of total suspended particles (TSP), size-fractionated particulate matter (PM\textsubscript{10}, PM\textsubscript{10-2.5}, PM\textsubscript{2.5}), sulfur dioxide (SO\textsubscript{2}) and oxides of nitrogen (NO\textsubscript{x}) substantially extended the upper end of the pollution ranges of previous air pollution epidemiological studies conducted in North America and Europe (Qian et al., 2001, Ware et al., 1986; Dockery et al., 1989, 1996; Spengler et al., 1996; Ackermann-Liebrich et al., 1997; Braun-Fahrlander et al., 1997; Zemp et al., 1999; Peters et al., 1999a,b; Pikhart et al., 2000). In the four Chinese cities, the 1993–1996 4-year means of TSP, SO\textsubscript{2}, and NO\textsubscript{x} ranged from 198 to 659 μg/m\textsuperscript{3}, 14.6 to 331 μg/m\textsuperscript{3}, and 31.5 to 239 μg/m\textsuperscript{3}, respectively. The 1995–1996 2-year means of PM\textsubscript{10}, PM\textsubscript{10-2.5}, PM\textsubscript{2.5} ranged from 80.7 to 232 μg/m\textsuperscript{3}, 29.2 to 107 μg/m\textsuperscript{3}, and 51.5 to 142 μg/m\textsuperscript{3}, respectively. However, few studies have been conducted to systematically examine adverse impacts on children’s health.
respiratory systems from long-term exposure to high levels of multiple pollutants in Chinese cities. Thus, the extent to which exposure to ambient air pollutant mixtures affects children’s respiratory health, in a free-living and community-dwelling scenario, is not yet clearly understood.

We, therefore, launched the Four Chinese Cities Study on health effects of urban ambient air pollution on the respiratory system of children and adults in the four Chinese cities of Chongqing, Guangzhou, Lanzhou, and Wuhan in 1993. Within each city, one urban district and one suburban district were selected for the study. In this paper, we examined associations between long-term exposure to ambient air pollutant mixtures and children’s respiratory symptom and illness prevalence in the four Chinese cities. The first step of our exploratory analysis involved the use of a hierarchical clustering technique to classify study districts into district clusters based on average ambient concentrations of TSP, PM$_{10-2.5}$, PM$_{2.5}$, SO$_2$, and NO$_x$. In the second step, we used these district clusters as new ecological variables in unconditional logistic regression models to estimate the respiratory conditions’ odds ratios with respect to district clusters. In the final step, we examine exposure–response associations by relating district-cluster-specific levels of individual pollutant concentrations to the estimated odds ratios. Since the district clusters were developed based on the “combined” levels of all of the measured pollutants, the above approach may shed light on examining the health effects of exposure to ambient air pollutant mixtures.

2. Methods

2.1. Data collection

The design and data collection of the Four Chinese Cities Study have been described in previous papers (Wei et al., 1999). Briefly, the Four Chinese Cities Study was designed to study respiratory health effects of long-term exposure to ambient (outdoor) and indoor air pollution in children and adults residing in the four cities. Ambient air pollutant measurements and questionnaire surveys were conducted from 1993 to 1996. In each city, an urban district and a suburban district were selected where TSP, PM$_{2.5}$, PM$_{10-2.5}$, SO$_2$, and NO$_x$ were measured (in total, eight districts were selected). In each district, all students from one or two elementary schools were recruited to participate in a questionnaire survey, and a subset of the students was selected for longitudinal lung function measurements. The lung function data are being processed, and the results will be reported elsewhere.

The questionnaire survey was described elsewhere (Qian et al., 2002). Before the questionnaire survey was conducted, written signed consents had been received from the students’ parents or guardians. In total, 7754 completed questionnaires were returned during 1993 to 1996, reflecting a very high return rate (Table 1). Some of the returned questionnaires were excluded from further analyses based on the reasons described below: (1) if a child had resided in his/her community for less than 2 years; (2) if a child’s household used wood as the main heating or cooking fuel (the percentage rates of wood heating or wood cooking were very small in seven of the eight districts (<5%) and the inclusion of household wood use may complicate the investigation of our major interest, coal smoke effects); and (3) if there were unanswered or ambiguously answered key questions that could not be clarified in the follow-up visits. After the screenings had been completed, a total of 7058 children remained in the subject pool. District-specific subject sample sizes ranged from 405 in the Chongqing suburban district to 1837 in the Wuhan urban district (Table 1).

2.2. Air pollution exposure assessment

Ambient air concentrations of TSP, SO$_2$, and NO$_x$ from 1993 to 1996 were obtained from the municipal
ambient air pollution monitoring stations in each of the eight study districts. The distance between the monitoring station and a selected elementary school or a subject residence was less than 8 km in each city under study. Concentrations of PM$_{2.5}$, PM$_{10-2.5}$ (and PM$_{10}$ = PM$_{2.5}$ + PM$_{10-2.5}$), as well as TSP, were also measured in the schoolyards of all the elementary schools for 2 years in 1995 to 1996. Based on these measurements of ambient concentrations during two different time periods, a cluster analysis technique was used to classify the eight districts into district clusters (Afifi and Clark, 1984; SAS Institute Inc., 1990; Qian et al., 2001). This is a reasonable procedure since the eight districts were not necessarily all unique in terms of ambient pollution levels. We performed the cluster analysis using the 2-year mean of PM$_{2.5}$, the 2-year mean of PM$_{10-2.5}$, the difference of the 2-year mean of TSP and PM$_{10}$ (TSP–PM$_{10}$), the 4-year mean of SO$_2$, and the 4-year mean of NO$_x$ in the eight study districts. The details of the cluster analysis and results were reported in the previous paper (Qian et al., 2001). Ambient exposure classification was thus made by district cluster with a fundamental assumption that all of the subjects under study living within a district cluster were exposed to a similar combined level of all of the measured pollutants (Lebowitz et al., 1975; Love et al., 1981; Chapman et al., 1985; Lioy, 1990; Von Mutius and Fritsch, 1992; Lioy and Pellizzari, 1995; Pope et al., 1995).

Exposure to some residential indoor air pollution sources was estimated based on the questionnaire responses to a set of relevant questions. In the current analysis, we constructed variables to represent exposure to parental smoking, exposure to heating coal smoke, and exposure to cooking coal smoke. Since very few mothers (<2.5%) in the four cities were smokers, father’s smoking status was used to represent children’s exposure to environmental tobacco smoking (ETS) at home. The estimation of exposure to heating coal smoke and cooking coal smoke was reported in detail elsewhere (Qian et al., 2002). Briefly, the estimation was made using a Scenario Evaluation Approach considering the total number of days of using coal for home heating and the total number of meals cooked using coal during the entire lifetime of a given child. Home ventilation conditions were also considered in a semi-quantitative manner. The subjects were then classified into four ordinal categories of exposure to cooking coal smoke, and four categories of exposure to heating coal smoke, as follows: heavily exposed, moderately exposed, lightly exposed, and control (no reported exposure). We constructed zero–one dummy variables to represent each exposure level other than the control level (totaling six variables). These variables, along with other covariates, were included in the statistical models. The other covariates considered included age, gender, whether sleeping in own room, mother’s education level, house type, and cooking oil type.

### 2.3. Definition of respiratory health outcomes

Six respiratory symptoms and illnesses were selected as health outcomes in this particular analysis. These outcomes were defined as follows based on the questionnaire responses: (1) Cough: the answers to several cough-related questions indicate that the study child has coughed for at least 1 month per year either with or apart from colds. (2) Phlegm: the answers to several phlegm-related questions indicate that the study child has brought up phlegm or mucus from the chest for at least 1 month per year either with or apart from colds. (3) Cough with phlegm: the answers indicate that the study child usually has a cough either with or apart from colds and that the child usually brings up phlegm or mucus from the chest either with or apart from colds. (4) Wheeze: the answers to wheeze-related questions indicate that the study child’s chest has at some time sounded wheezy or whistling either with or apart from colds. (5) Asthma: a yes answer to the question “Has a doctor ever diagnosed asthma in this child?” (6) Bronchitis: a yes answer to the question “Has a doctor ever diagnosed bronchitis in this child?”

### 2.4. Statistical methods

The cluster analysis was reported elsewhere (Qian et al., 2001). Briefly, each of the eight study districts had five measurements [the mean concentrations of PM$_{2.5}$, PM$_{10-2.5}$, (TSP–PM$_{10}$), SO$_2$, and NO$_x$] and each of the eight districts may be considered a point in five-dimensional Euclidean space. The closeness or similarity between any of the two points was expressed by their Euclidian distance (the square root of the sum of the squared differences between the coordinates of the two points) (Afifi and Clark, 1984). In addition to this definition of the distance metric, we also defined “single linkage” as a rule to guide linkages (Vogt et al., 1987; SAS Institute Inc., 1990; Howard, 1991). We began with eight clusters; i.e., each point constituted its own cluster. We combined the two closest clusters as a group in successive steps, thus reducing the number of clusters by one in each step. In the final step, we grouped all points into one cluster. We determined the number of clusters by checking the minimum distance between clusters at each successive step, i.e., we chose a cluster when the minimum distance exceeded the specific value of 0.12 (Afifi and Clark, 1984).

Unconditional logistic regression models were used to calculate covariates-adjusted odds ratios (ORs) of each health outcome with respect to ecological variables (i.e., district clusters). The covariates (independent variables) were age, gender, whether sleeping in own room, mother’s education level, house type, cooking oil type, heating coal smoke, cooking coal smoke, and father’s smoking status. The models were constructed with only main effects and then with main effects and interaction terms. The tested interaction terms were: (1) district cluster by heating coal...
Table 2
Arithmetic mean concentrations of ambient air pollutants in the four district clusters (ug/m³), 1993–1996

<table>
<thead>
<tr>
<th>Pollutants</th>
<th>Cluster 1b</th>
<th>Cluster 2c</th>
<th>Cluster 3d</th>
<th>Cluster 4e</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSP–PM_{10}</td>
<td>101</td>
<td>116</td>
<td>258</td>
<td>437</td>
</tr>
<tr>
<td>PM_{2.5–10}</td>
<td>39</td>
<td>59</td>
<td>60</td>
<td>107</td>
</tr>
<tr>
<td>PM_{2.5}</td>
<td>61</td>
<td>99</td>
<td>96</td>
<td>115</td>
</tr>
<tr>
<td>SO_{2}</td>
<td>28</td>
<td>92</td>
<td>193</td>
<td>130</td>
</tr>
<tr>
<td>NO_{3}</td>
<td>55</td>
<td>131</td>
<td>68</td>
<td>77</td>
</tr>
</tbody>
</table>

a The four district clusters came from the cluster analysis of PM_{2.5}, PM_{10–2.5}, (TSP–PM_{10}), SO_{2}, and NO_{3}. Two-year mean concentrations are presented here for PM_{2.5}, PM_{2.5–10}, and PM_{10} during 1995 to 1996.
b Cluster 1 includes the Guangzhou and Wuhan suburban districts.
c Cluster 2 includes the Guangzhou and Wuhan urban districts, and the Chongqing suburban district.
d Cluster 3 includes the Chongqing urban and Lanzhou suburban districts.
e Cluster 4 includes the Lanzhou urban district.

The eight study districts were grouped into four district clusters based on the combined levels of all the pollutants used in the analysis (Qian et al., 2001). Mean concentrations of the pollutants included in the cluster analysis are shown in Table 2 by district cluster. Cluster 1 consists of the two suburban districts in Guangzhou and Wuhan; this district cluster appeared to have the lowest combined pollution level of PM_{2.5}, PM_{10–2.5}, (TSP–PM_{10}), SO_{2}, and NO_{3}. Concentrations of the pollutants included in the cluster analysis were used in all of the analyses (Breslow and Day, 1980; SAS Institute Inc., 1995). We found that the main-effects-only models fit the data better than the main-effects-and-interactions models, on the basis of the Akaike Information Criterion and the Schwartz Criterion (SAS Institute Inc., 1995). Thus, only the results of the main effects models are reported here.

3. Results

3.1. Ambient concentrations of district cluster

The eight study districts were grouped into four district clusters based on the combined levels of all the pollutants used in the analysis (Qian et al., 2001). Mean concentrations of the pollutants included in the cluster analysis are shown in Table 2 by district cluster. Cluster 1 consists of the two suburban districts in Guangzhou and Wuhan; this district cluster appeared to have the lowest combined pollution level of PM_{2.5}, PM_{10–2.5}, (TSP–PM_{10}), SO_{2}, and NO_{3}. Concentrations of the pollutants included in the cluster analysis were used in all of the analyses (Breslow and Day, 1980; SAS Institute Inc., 1995). We found that the main-effects-only models fit the data better than the main-effects-and-interactions models, on the basis of the Akaike Information Criterion and the Schwartz Criterion (SAS Institute Inc., 1995). Thus, only the results of the main effects models are reported here.

A closer look at Table 2 shows that, overall, the district classification was primarily driven by the particulate matter pollution, especially by the TSP–PM_{10}. The TSP–PM_{10} and PM_{10–2.5} levels had a gradient across the four district clusters in the following order: Cluster 4>Cluster 3>Cluster 2>Cluster 1, the same as the ranking for the combined pollution level. The TSP–PM_{10} concentration in Cluster 4 was about 4.3 times higher than that observed in Cluster 1. The difference between clusters for PM_{10–2.5} and PM_{2.5} were less profound than for TSP–PM_{10}; although the concentrations in Cluster 4 were also the highest for PM_{10–2.5} and PM_{2.5}. The gradients of PM_{2.5}, SO_{2}, and NO_{3} were different from that of TSP–PM_{10}. Cluster 4>Cluster 2>Cluster 3>Cluster 1 for PM_{2.5}; Cluster 3>Cluster 4>Cluster 2>Cluster 1 for SO_{2}; and Cluster 2>Cluster

Table 3
Distributions of subject and household characteristics in the four district clusters

<table>
<thead>
<tr>
<th>Variables/categories</th>
<th>Cluster 1 (%)</th>
<th>Cluster 2 (%)</th>
<th>Cluster 3 (%)</th>
<th>Cluster 4 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–9</td>
<td>762</td>
<td>61.5</td>
<td>1655</td>
<td>67.1</td>
</tr>
<tr>
<td>10–16</td>
<td>478</td>
<td>38.5</td>
<td>812</td>
<td>32.9</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>608</td>
<td>49.0</td>
<td>1215</td>
<td>49.3</td>
</tr>
<tr>
<td>Female</td>
<td>632</td>
<td>51.0</td>
<td>1252</td>
<td>50.7</td>
</tr>
<tr>
<td>Mother’s education level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal to or lower than middle school</td>
<td>733</td>
<td>59.1</td>
<td>1882</td>
<td>76.3</td>
</tr>
<tr>
<td>Higher than middle school</td>
<td>507</td>
<td>40.9</td>
<td>585</td>
<td>3.7</td>
</tr>
<tr>
<td>Paternal smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>813</td>
<td>65.6</td>
<td>1991</td>
<td>80.7</td>
</tr>
<tr>
<td>Yes</td>
<td>427</td>
<td>34.4</td>
<td>476</td>
<td>19.3</td>
</tr>
<tr>
<td>Sleeping in own room</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>547</td>
<td>44.1</td>
<td>693</td>
<td>28.1</td>
</tr>
<tr>
<td>No</td>
<td>693</td>
<td>55.9</td>
<td>1774</td>
<td>71.9</td>
</tr>
<tr>
<td>Cooking oil type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapeseed oil</td>
<td>256</td>
<td>20.7</td>
<td>920</td>
<td>7.3</td>
</tr>
<tr>
<td>Other oils</td>
<td>984</td>
<td>79.4</td>
<td>1547</td>
<td>62.7</td>
</tr>
<tr>
<td>Heating coal smoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (0 days)</td>
<td>1045</td>
<td>84.3</td>
<td>1871</td>
<td>75.8</td>
</tr>
<tr>
<td>Lightly exposed (1–465 days)</td>
<td>99</td>
<td>8.0</td>
<td>310</td>
<td>12.6</td>
</tr>
<tr>
<td>Moderately exposed (466–630 days)</td>
<td>41</td>
<td>3.3</td>
<td>138</td>
<td>5.6</td>
</tr>
<tr>
<td>Heavily exposed (631–2310 days)</td>
<td>55</td>
<td>11.6</td>
<td>448</td>
<td>19.9</td>
</tr>
<tr>
<td>Cooking coal smoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (0 meals)</td>
<td>533</td>
<td>43.0</td>
<td>1138</td>
<td>46.1</td>
</tr>
<tr>
<td>Lightly exposed (0–387 meals)</td>
<td>206</td>
<td>16.6</td>
<td>277</td>
<td>11.2</td>
</tr>
<tr>
<td>Moderately exposed (388–930 meals)</td>
<td>209</td>
<td>16.9</td>
<td>432</td>
<td>17.5</td>
</tr>
<tr>
<td>Heavily exposed (931–10816 meals)</td>
<td>292</td>
<td>23.6</td>
<td>620</td>
<td>25.1</td>
</tr>
</tbody>
</table>

See Table 2 for the definitions of Clusters 1, 2, 3, and 4.
4>Cluster 3>Cluster 1 for NO$_x$. These different ranking orders for different pollutants were used to examine prevalence–concentration relationships at individual pollutant levels, as shown later in the paper.

3.2. Distributions of covariates

The distributions of subject and household characteristics are presented in Table 3. The subjects had an age range of 5 to 16 years old. The overall male/female ratio was approximately 1. A majority of the children’s mothers had received low levels of education not exceeding middle school (no school, attended or graduated from primary school, attended or graduated from middle school) in the four district clusters. Given that very few mothers (<2.5%) were smokers, only paternal smoking status was included in the analysis. Paternal smoking status was defined as “yes” when a father had smoked more than 100 cigarettes or an equivalent amount of tobacco in his entire life. On the basis of this definition, the percentages of smoking fathers ranged from 65.6 in Cluster 1 to 80.7 in Cluster 2. The majority of children in the study had shared a room with other family members. Cooking with rapeseed oil was considered in the analyses because this cooking oil had been one of the most commonly used cooking oils in the four Chinese cities. Estimated lifetime exposures to heating coal smoke and cooking coal smoke, by a Scenario Evaluation Approach, were grouped into four ordinal categories, control (no reported exposure), lightly exposed, moderately exposed, and heavily exposed (Qian et al., 2002).

3.3. Respiratory health outcomes by district clusters

Cluster-specific crude prevalence rates of the six respiratory health outcomes are shown in Table 4. Chi-square tests indicated that the crude prevalence rates of cough, phlegm, cough with phlegm, bronchitis, wheeze, and asthma were significantly different across the four district clusters ($p<0.01$) with the highest rates in Cluster 4 and the lowest in Cluster 1. Interestingly, the crude prevalence rates of phlegm, cough with phlegm, bronchitis, and wheeze had the same ranking order of the combined pollution levels (Cluster 4>Cluster 3>Cluster 2>Cluster 1). The crude prevalence rates of cough and asthma, however, were higher for Cluster 2 than for Cluster 3.

After adjustment for all of the covariates mentioned earlier, estimated odds ratios (ORs) and 95% confidence intervals (CIs), with respect to the four district clusters, are shown in Table 5. Cluster 1 was regarded as the reference level (OR = 1) since all of the crude prevalence rates were the lowest in this cluster (Table 4). We observed that the ORs of cough with phlegm and wheeze had the same ranking order of the combined pollution levels (Cluster 4>Cluster 3>Cluster 2>Cluster 1) and the OR values for the other three clusters were significantly greater compared to the reference cluster (Cluster 1). The ORs of phlegm were significantly higher for the three clusters than the reference, but the OR was not greater for Cluster 3 than for Cluster 2. The ORs of cough and asthma were significantly higher for Cluster 2 and Cluster 4 than the reference, but not for Cluster 3. The OR of bronchitis was significantly higher only for Cluster 4.

4. Discussion

By relating concentration levels in each district cluster and OR and 95% CI values (Tables 2 and 5), we observed monotonic, positive, and statistically significant concentration–symptom relationships between ambient air pollutant mixture exposure and prevalence of cough with phlegm and

<table>
<thead>
<tr>
<th>Respiratory outcomes</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Cough</td>
<td>1.00</td>
<td>1.53 1.22</td>
<td>1.92 1.12</td>
<td>0.83 1.53</td>
</tr>
<tr>
<td>Phlegm</td>
<td>1.00</td>
<td>1.77 1.21</td>
<td>2.60 1.76</td>
<td>1.10 2.81</td>
</tr>
<tr>
<td>Cough with phlegm</td>
<td>1.00</td>
<td>1.86 1.53</td>
<td>2.24 2.06</td>
<td>1.63 2.61</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>1.00</td>
<td>1.13 0.97</td>
<td>1.31 1.03</td>
<td>0.84 1.26</td>
</tr>
<tr>
<td>Wheeze</td>
<td>1.00</td>
<td>1.72 1.39</td>
<td>2.14 1.84</td>
<td>1.40 2.41</td>
</tr>
<tr>
<td>Asthma</td>
<td>1.00</td>
<td>1.55 1.01</td>
<td>2.38 1.30</td>
<td>0.74 2.27</td>
</tr>
</tbody>
</table>

See Table 2 for the definitions of Clusters 1, 2, 3, and 4.

Table 4
Crude prevalence rates of respiratory symptoms and illnesses, by district cluster

Table 5
Odds ratios and 95% confidence intervals for children’s respiratory symptoms and illnesses with respect to district clusters

See Table 2 for the definitions of Clusters 1, 2, 3, and 4.

4>Cluster 3>Cluster 1 for NO$_x$. These different ranking orders for different pollutants were used to examine prevalence–concentration relationships at individual pollutant levels, as shown later in the paper.
wheeze. This suggests adverse effects of the pollutant mixture on children’s respiratory health, specifically in increasing the prevalence of cough with phlegm and wheeze. Because Clusters 1–3 contains districts from different cities and Cluster 4 only contains the Lanzhou urban district, the observed effects were less likely to be driven by differences in other aspects (e.g., climate, population) across the different cities. This is one of the benefits of using the cluster analysis method.

The observed monotonic, positive exposure–response relationships were driven by PM pollution levels across the four district clusters. These observations are consistent with those from our previous studies on different populations, which demonstrated positive associations of ambient PM (TSP) exposure with prevalence of cough and phlegm across three Chinese cities of Guangzhou, Wuhan, and Lanzhou (Qian et al., 2000). In addition, these observations are also generally consistent with those from other Chinese studies. He et al. (1993) found that the children’s exposure to higher TSP consistently decreased the lung function for children living in urban area of Wuhan. This decreased pulmonary function was generally consistent with the results of upper respiratory tract examinations in that the prevalence rate of upper respiratory irritations was generally elevated in urban children, compared to the suburban children in Wuhan. Xu et al. (2000) also observed a significant association of exposure to TSP with a high daily mortality in Shenyang of China. Therefore, the overall consistency of evidence between the present study, our previous findings, and other Chinese studies may suggest a coherence of health effects of children’s long-term exposures to high PM levels in Chinese cities (Bates, 1992).

Evidence has been documented that fine PM is more strongly associated with certain health outcomes (e.g., mortality rate, cardiovascular morbidity) than coarse PM (PM$_{10-2.5}$). Our previous pilot studies and other Chinese studies could not separate health effects of exposure to a single PM fraction such as PM$_{2.5}$ since there is no historical data on size fractionated PM measurements (Zhang et al., 1999). One concern would then arise that the observed TSP health effects may be the ones associated with particles of certain size rather than with TSP. With these considerations

![Fig. 1. Associations between arithmetic mean concentrations of ambient PM$_{2.5}$ and estimated ORs of cough, phlegm, bronchitis, and asthma across the four district clusters.](image-url)
we examine exposure–response relationships by relating district-cluster-specific levels of individual size fractionated PM concentrations to the estimated odds ratios. We observed that the ORs of cough, phlegm, bronchitis, and asthma had the same ranking order of the PM\textsubscript{2.5} concentrations (Cluster 4>Cluster 3>Cluster 2>Cluster 1, Fig. 1); and the ORs of cough with phlegm and wheeze had the same ranking order of the PM\textsubscript{10–2.5} concentrations as well as TSP concentrations (Figs. 2 and 3). These observations may indicate that different sizes of particles at the high pollution levels such as occurred in this study were independently associated with respiratory health outcomes under study.

In China, environmental standards for PM have not been strictly set up and executed. For example, the Chinese 24-h TSP standard for residential areas is 300 µg/m\(^3\), the newly promulgated daily PM\textsubscript{10} standard is 150 µg/m\(^3\), and there is no PM\textsubscript{2.5} standard. Our previous paper indicates that 7% to 60% of the daily TSP measurements exceeded 300 µg/m\(^3\); 6% to 70% of the daily PM\textsubscript{10} measurements exceeded 150 µg/m\(^3\); and 26% to 79% of the daily PM\textsubscript{2.5} measurements exceeded 65 µg/m\(^3\) (the U.S. daily PM\textsubscript{2.5} standard) across the four Chinese cities (Qian et al., 2001). On the basis of this evidence, it would be an effective intervention to improve children’s respiratory health by launching a broad-based campaign to abate ambient airborne particulate pollution levels in China.

Air pollution epidemiological research has tended to focus on single pollutants’ health effects because of the...
regulatory focus on specific compounds (e.g., SO$_2$) or groups of compounds (Dockery et al., 1989; Raizenne et al., 1996; Ackermann-Liebrich et al., 1997; Peters et al., 1999a). Our literature review indicates that published studies have generally reported relative health effects of exposure to a single pollutant or two pollutants by using one or two pollutant models, and we are still limited in terms of statistical approaches to assess exposure to ambient air pollutant mixtures (Schwartz, 1989; Dockery et al., 1996; Peters et al., 1999b; Jedrychowski et al., 1999; Pikhart et al., 2000). People are simultaneously exposed to multiple ambient air pollutants in a real life scenario. The health effects, indexed by the commonly used health outcomes including the six respiratory conditions in the present study, are not specifically attributable to single pollutants, but most likely to be attributable to pollutant mixtures. Thus, it is worthwhile making new efforts to assess the exposure to pollutant mixtures and their health effects. In this study, we explored the application of a cluster analysis in assessing exposure to ambient air pollutant mixtures and examined the associations between the exposure and the six respiratory conditions. The approach taken here may provide a supplemental method to develop an ecological indicator variable used in statistical models to examine health effects of exposure to ambient pollutant mixtures. When ambient multiple pollutants are strongly correlated as shown in our previous report (Qian et al., 2001), this approach may be preferable, since correlated ambient pollution levels do not complicate this type of analysis (Vogt et al., 1987; SAS Institute Inc., 1990; Howard, 1991).

However, there are several methodological concerns that needed to be discussed here. First, the major limitation of this approach is that the toxicological importance of different pollutants cannot be reasonably weighted because of an apparent lack of any biologically demonstrable mechanism (Lison et al., 1997; Osier and Oberdorster, 1997; Oberdorster et al., 1994; Li et al., 1999). Second, the cluster analysis approach is highly empirical and may have significant shortcomings in grouping the districts into the district clusters (Kachigan, 1991; Everitt, 1993; Cody and Smith, 1997). For example, the approach could not evaluate the extent that urban–suburban differences contributed to the observed effect since the observed health effects of the exposure to ambient air pollutant mixtures were obtained based on the comparison of ORs in the four district clusters rather than in the eight districts. Districts within a district cluster may differ with respect to their characteristics, other known factors, as well as air pollution. Based on this consideration, our previous data analyses were also performed at district levels, i.e., indicator variables of the eight study districts instead of the four district clusters were included in the logistic regression models. The results are generally in agreement with those of the present study, but the observed health effects are more pronounced in the present study (Qian, 2001). These observations may indicate that the estimated risks were more likely to be associated with exposure to multiple ambient air pollutants, rather than exposure to a single pollutant, since the exposure to the pollutant mixture in the district cluster levels may be “more similar” in terms of causing respiratory health effects in children. Third, this study relies largely upon self-reported respiratory symptoms for the majority of health outcomes. The potential for misclassification and/or recall bias is inevitable like all epidemiological research (Hennekens and Buring, 1987). However, our questionnaire survey was accomplished through careful design and meticulous conduct of the study as detailed shown as below. All of the investigators in the four cities were centrally and rigorously trained using the standardized questionnaire, which contains highly objective, uniform criteria, and closed-ended questions to collect the information on the respiratory symptoms. We minimized the need for interpretation on the part of both the investigators and the parents or guardians. Both the investigators and the parents or guardians maintained blindness to the greatest extent possible in terms of specific hypotheses under investigation and the study children’s relative exposure status across the original eight study districts. All of the above makes it likely that the responses were nondifferential in the eight study districts (Kelsey et al., 1996). These nondifferential responses would lead to underestimating the associations (Rothman and Greenland, 1998). Last, in the present study, the observations on associations between ambient air pollutant mixture exposure and respiratory health outcomes were made based on a fundamental assumption that community (district cluster) mean concentrations of ambient air PM$_{2.5}$, PM$_{2.5–10}$, PM$_{10}$, TSP, SO$_2$, and NO$_x$ serve as surrogates for personal exposure as well as dosage. However, there may be a large uncertainty in the relationship between ambient air concentrations and personal exposure (Quackenboss et al., 1989a,b; Wilson, 1996; Zhang et al., 1999; Peters et al., 1999a,b; Hrubu et al., 2001). Nevertheless, since the high levels of PM$_{2.5}$, PM$_{2.5–10}$, PM$_{10}$, TSP, SO$_2$, and NO$_x$ existed for a long time in the study districts, and none of the residences in the study districts were built with energy conservation in mind, these high levels of the pollutants could easily infiltrate into the homes (Huang et al., 1990; Zhou et al., 1990; Waldman et al., 1991; Chen et al., 1992; He et al., 1993; Wei et al., 1999). The outdoor pollution may then be dominant, making indoor pollution emission sources less significant. Thus, the uncertainties brought by using the community mean concentrations may not be significant. However, this explanation needs to be further verified because no direct measurements of indoor air quality were made from the present study.

Acknowledgements

The authors are grateful for the essential contribution made by members of the field study staff at Chongqing Environmental Monitoring Station, Guanghzou Environ-
References


