Archives of Environmental & Occupational Health, Vol. 68, No. 4, 2013
Published with license by Taylor & Francis Group, LLC

Respiratory Disease in Relation to Outdoor Air Pollution in Kanpur, India

Hai-Ying Liu, PhD; Alena Bartonova, RNDr; Martin Schindler, PhD; Mukesh Sharma, PhD; Sailesh N. Behera, PhD; Kamlesh Katiyar, PhD; Onkar Dikshit, PhD

ABSTRACT. This paper examines the effect of outdoor air pollution on respiratory disease in Kanpur, India, based on data from 2006. Exposure to air pollution is represented by annual emissions of sulfur dioxide (SO_2), particulate matter (PM), and nitrogen oxides (NO_x) from 11 source categories, established as a geographic information system (GIS)-based emission inventory in 2 km × 2 km grid. Respiratory disease is represented by number of patients who visited specialist pulmonary hospital with symptoms of respiratory disease. The results showed that (1) the main sources of air pollution are industries, domestic fuel burning, and vehicles; (2) the emissions of PM per grid are strongly correlated to the emissions of SO_2 and NO_x; and (3) there is a strong correlation between visits to a hospital due to respiratory disease and emission strength in the area of residence. These results clearly indicate that appropriate health and environmental monitoring, actions to reduce emissions to air, and further studies that would allow assessing the development in health status are necessary.

[Supplementary materials are available for this article. Go to the publisher's online edition of *Archives of Environmental & Occupational Health* for material on emission of SO_2, PM, NO_x from various sources, and total number of inhabitants, total number of patients in grid squares covering the Kanpur city.]

KEYWORDS: emission, outdoor air pollution, patients, respiratory disease

It is well known that air pollution causes respiratory and cardiovascular diseases.[1–3] Economic development, urbanization, energy consumption, transportation, and rapid population growth are major driving forces of air pollution in large cities, especially in megacities.[4] Air pollution levels in developed countries have been decreasing dramatically in recent decades. However, in other countries, air pollution levels are still relatively high, although the levels have been gradually decreasing or have remained stable despite rapid economic development.[4,5] In recent years, several hundred epidemiological studies, time-series studies in particular, have been conducted in developed countries, on short- and long-term effects of air pollution on human health covering different age groups, including children and young and old adults.[6–8] Research has shown that long-term exposure to air pollutants increases the risk of respiratory illnesses such as allergies, asthma, chronic obstructive pulmonary disease, and lung cancer.[9,10] Children and elderly persons are particularly vulnerable to health effects of ozone (O_3), particulate matter (PM), and other airborne toxicants.[11,12] Hospital admissions have been recognized as a more sensitive marker than mortality for assessment of the air pollution effects on human health.[13,14] A literature review on outdoor air pollution and health in Asia identified over 400 studies of health effects of air pollution in 13 countries during the period 1980 to 2007.[15] Over 80 time-series studies conducted in Asian cities also showed similar spectrum of adverse health effects from acute and chronic respiratory symptoms and changes in pulmonary function to increased mortality from cardiovascular or respiratory diseases or lung cancer, associated

© Hai-Ying Liu, Alena Bartonova, Martin Schindler, Mukesh Sharma, Sailesh N. Behera, Kamlesh Katiyar, and Onkar Dikshit

Hai-Ying Liu *and* **Alena Bartonova** *are with the Norwegian Institute for Air Research (NILU), Kjeller, Norway*. **Martin Schindler** *is with the Department of Applied Mathematics, Technical University of Liberec, Czech Republic*. **Mukesh Sharma**, **Sailesh N. Behera**, *and* **Onkar Dikshit** *are with the Department of Civil Engineer, Indian Institute of Technology Kanpur, Uttar Pradesh, India*. **Kamlesh Katiyar** *is with Ganesh Shankar Vidyarthi Memorial (GSVM) Medical College, Swaroop Nagar, Kanpur, Uttar Pradesh, India*.

with exposure to PM, sulfur dioxide (SO_2), nitrogen dioxide (NO_2) and O_3, to those explored in Europe and North America.[15–17]

Air pollution poses a global challenge, as adverse effects still exist even at relatively low air pollutant concentrations, and there may not be any safe or threshold pollution levels. The air pollution problem is more severe in Asian countries due to high pollution levels and high population densities.[4,15–17] Important gaps still remain, as the studies do not necessarily cover enough variations in urban settlements, pollution levels, and economic conditions.[15]

Many of the Indian cities are highly polluted and available mortality and morbidity statistics indicate that respiratory infections and chronic conditions are widespread.[18–24] However, there are very few Indian studies that link air pollution exposure to lung function deterioration, underlining the need for systematic studies.[21]

We carried out a pilot study to examine the correlation between outdoor air pollution (represented as annual emissions) and hospital visits due to respiratory disease in Kanpur (latitude 26°45′ north and longitude 80°15′ east; Figure 1), one of the most populous (4 million persons), industrialized, and polluted cities in India.[18,22,25]

Pollution loads/emissions were estimated for SO_2, PM, and nitrogen oxides (NO_x), whereas respiratory diseases were estimated from total number of registered respiratory patients visits to the largest chest and tuberculosis hospital (Lala Lajpat Rai [LLR] Hospital) in Kanpur, Uttar Pradesh, India.[22,26] This hospital receives patients from the whole district.

The broad objective of this study was to develop a methodology for air pollution health impact assessment, and to apply this approach in the city of Kanpur. The specific objective was to assess population-wide health effects of air pollution. This study lays a basis for further environmental health assessments, and provides rationale for improved air quality monitoring as well as mitigation measures.

METHODS

Data collection

Sampling Grids for Emission Assessment

The entire 28 km × 22 km area including the city was divided into 154 grids of equal size of 2 km × 2 km. Coordinates of the center point in each grid were recorded by the

Fig. 1. Location of Kanpur city in India (color figure available online).

Table 1.—Overview of Grid Squares for Detailed Emission Survey

Center point number	Stations	Latitude	Longitude	Lands use types
1	Sidbi Centre	26°30′36″	80°13′48″	Institutional
2	Vikas Nagar	26°29′24″	80°17′24″	Residential
3	Dada Nagar	26°28′12″	80°20′24″	Commercial
4	Colonel Ganj	26°28′12″	80°20′24″	Commercial
5	Pared	26°27′00″	80°17′24″	Industrial
6	Ramadevi	26°26′24″	80°19′12″	Residential
7	Juhilal Colony	26°24′36″	80°23′24″	Residential

Global Positioning System (GPS) in Universal Transverse Mercator (UTM) system. The other recorded items include the grid number and land use type. Seven key grids inside the city (Table 1), having varied land use types, were studied in detail by conducting house to house surveys of each polluting activity in the grid (Figure 2).

Pollution Sources

Pollution sources taken into account in the emission inventory include vehicles, domestic fuel burning, garbage burning, restaurants, diesel generators sets, medical-waste incinerators, funeral pyre burning, industries with stack height lower than 25 m (area sources), industries with stack height higher than 25 m (point sources), soil-road dust, construction and demolition recycling (Supplement Files 1–3). For point sources, the recorded items include the point sources number, grid number, industry name, production capacity (t/day), fuel consumption (t/day, fuel type), stack height (m), and latitude and longitude (Figure 2, Table 2).

Emissions as Surrogate Measure of Exposure

Each of the 154 grids was assigned a land use type based on the landscape map from the Central Pollution Control Board (CPCB), Delhi, India. Six land use types were identified (ie, agricultural, commercial, industrial, institutional, protected, and residential) in the entire city (Figure 3).

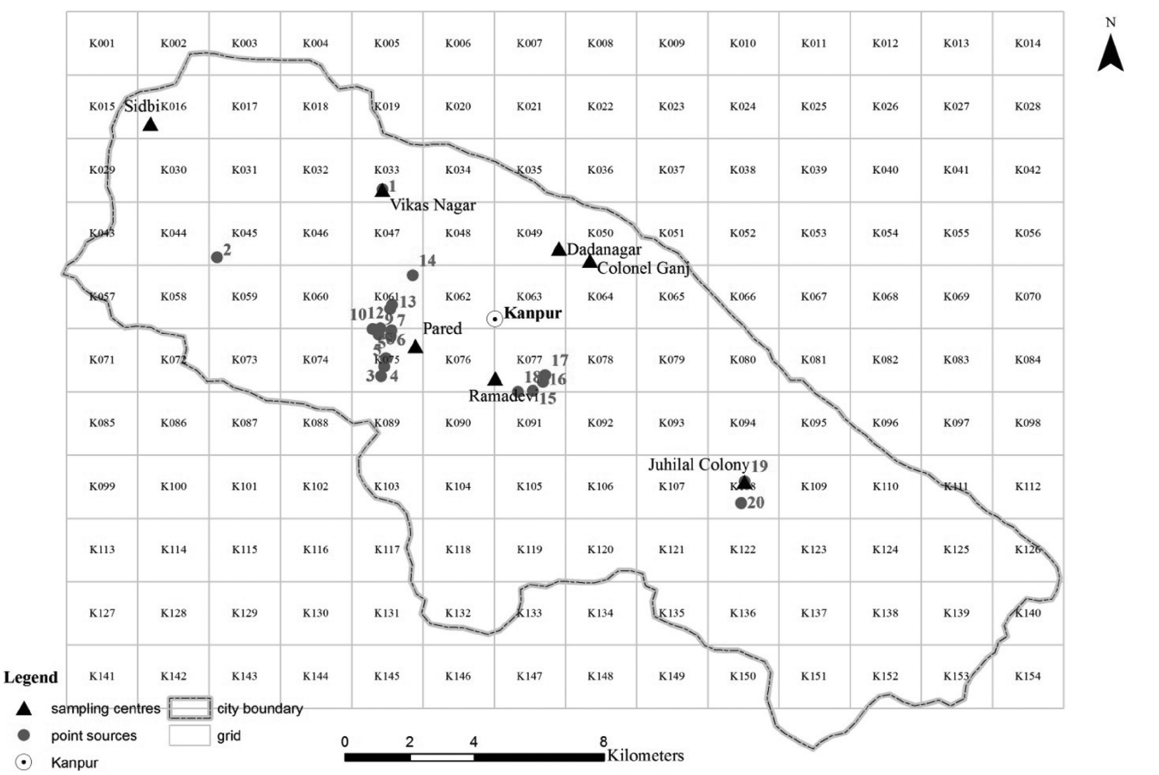

Fig. 2. Grid position over Kanpur, with the 7 grid squares with detailed emission survey (solid black triangle; Table 1), and industries' point sources (Table 2) (color figure available online).

Table 2.—Overview of Industrial Point Sources

Point sources number	Grid number	Industries	Production capacity (t/day)	Fuel consumption (t/day, coal)	Stack height (m)	Latitude	Longitude
1	K033	Rice mill	48	0.64	30	26°29′40″	80°17′11″
2	K045	Thermal power plant	314	2,030	120	26°28′31″	80°14′26″
3	K075	Iron and steel industry	32	2.2	30	26°26′30″	80°17′10″
4	K075	Textile industry	7	700 L/day (diesel)	60	26°26′40″	80°17′13″
5	K075	Rice mill	150	2	35	26°26′48″	80°17′15″
6	K075	Iron and steel industry	20	20 L/day (diesel) and 18 t/day (coal)	25	26°27′11″	80°17′20″
7	K075	Iron and steel industry	50	4.5	30	26°27′12″	80°17′20″
8	K061–K075	Iron and steel industry	28	3	25	26°27′19″	80°17′10″
9	K061–K075	Oil industry	2	2	25	26°27′13″	80°17′8″
10	K061–K075	Textile industry	5	1	55	26°27′18″	80°17′2″
11	K075	Rice mill	25	0.375	30	26°27′16″	80°17′20″
12	K061	Textile industry	6	1	60	26°27′39″	80°17′19″
13	K061	Leather industry	5	960 L/day (diesel)	35	26°27′43″	80°17′21″
14	K061	Iron and steel industry	11	10 L/day (diesel) and 1 t/day (coal)	25	26°28′13″	80°17′42″
15	K077–K091	Oil industry	8	750 L/day (diesel)	30	26°26′15″	80°19′42″
16	K077	Rice mill	36	0.48	35	26°26′24″	80°19′52″
17	K077	Leather industry	5	5	35	26°26′31″	80°19′55″
18	K077–K091	Rubber industry	10	2	30	26°26′14″	80°19′27″
19	K108	Textile industry	12	1.2	80	26°24′43″	80°23′14″
20	K108	Rice mill	35	0.425	35	26°24′21″	80°23′11″

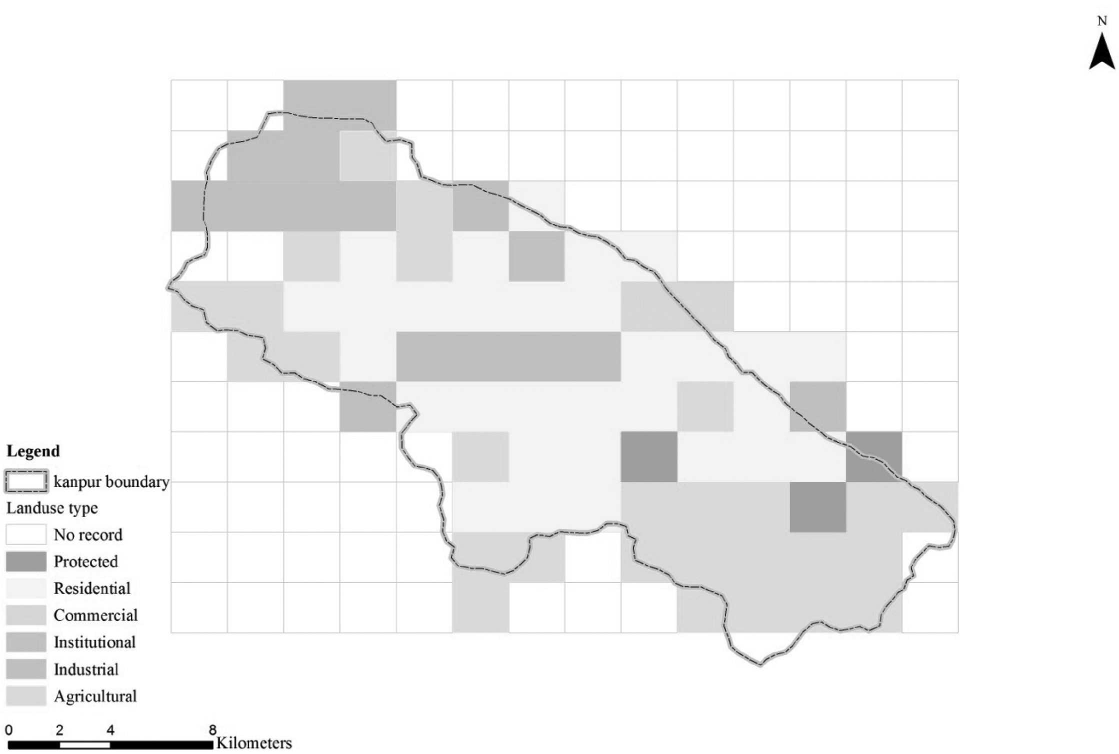

Fig. 3. **Grid squares and land use types (color figure available online).**

Based on the activity data of 7 grids with different land use, and 20 industries point pollution sources in detail (see Pollution Sources), the emissions in the other grids were obtained by mapping the 7 grids to the land use types of the other grids and by accounting for road length, number of vehicles on the road, and population in the grid. Emissions were estimated for PM, SO_2, and NO_x (for details of emission inventory, see Supplement Files 1, 2, 3).[26]

In Kanpur, several monitoring stations are in place, operated by state and central authorities. Their placement is appropriate for air quality assessment, but not for estimating exposure. They offer daily time series, but do not capture the variability of the outdoor pollution field. For this reason, the study team decided to use the emission estimates as a surrogate measure of exposure. In this way, the time resolution is lost, but we gain exposure gradients within the city, which seems appropriate in relation to the nature of the other information in this study.

Health Data

Health data were collected from the LLR Hospital records. The data include grid number (representing residence of patients, Supplement File 4), age, sex, smoking status, occupation, respiratory symptoms, and resident location/street name. The data on 8,557 patients who visited the hospital at least once were collected for the period of January 10 to December 29, 2006. Quality assurance procedure required 2 physicians recording the data from journals to electronic form, and periodic double entries of portion of the data. The data entry resulted in 8,340 valid records. Of these, 3,948 had identified the domicile (home address). The large number of symptoms on the medical record was classified into 12 categories (Table 3).

Table 3.—Respiratory Disease Symptom Categories and Total Number of Patients Diagnosed in Each Category

Serial number	Symptoms	Total number of patients
1	Abdominal pain and epigastric pain	391
2	Breathlessness and dyspnea	4,318
3	Chest pain	3,446
4	Common cold	440
5	Bough	6,433
6	Fever	4,356
7	Hemoptysis and similar symptom	929
8	Loss of appetite	1,249
9	Weakness	475
10	Other aches and pains	626
11	Swellings	375
12	Other unspecified symptoms	544

Data analysis

The analysis aimed to relate the data on hospital visits to exposure index represented by the emissions. There are 3 emission parameters, and we first investigated their variability and dependencies (see Emission Clusters). Then, we investigated the relationship between the emissions and the hospital visits (see Correlation Between Air Pollution and Respiratory Disease), and we attempted to analyze the relationship between emissions and respiratory symptoms (see Analysis of Occurrence for a Respiratory Disease Symptom). Finally, we have investigated the temporal pattern of the hospital visits (see Analysis of Seasonal Differences). R version 2.7.1,[27–29] packages "cluster" and "rgl"[30] and clustering functions "kmeans" and "pam," were used for data analysis. ArcGIS 9.0 (ESRI, Redlands, CA, USA) was used for all data creation and map illustration.

A cluster analysis[31] was conducted on the 3-dimensional emission data (SO_2, NO_x, PM) for each grid. The aim of this analysis is to group grids into certain number of clusters based on the level of air pollution load in the grids. We included emission data from all the pollution sources (see Pollution Sources), disregarding the point source from stacks higher than 25 m that are likely to influence larger area than the immediate grid. This step allowed us to conduct the analysis using statistical methods without incorporation of spatial dependence or autocorrelation, only using classical logistic regression model. We disregard also the observations in grids for which "land use" type is "protected." For clustering, we used R functions "kmeans" and "pam" (a robust version of "kmeans") from package "cluster." Within groups, "sum of squares" and "average silhouette plots" were used to determine the optimal number of clusters.

The relative number of hospital visits due to respiratory disease/symptoms was compared between the 4 clusters. In this analysis, the total number of inhabitants in each cluster was taken into account. The Pearson's chi-square test[32] for independence in the contingency table of number of respiratory patients and number of inhabitants who did not visit the hospital was conducted. Furthermore, we used a logistic regression[33–35] to model an occurrence of a patient visit in LLR Hospital in relation to the cluster to which the home grid square belongs. The people who did not visit the hospital are the control group.

Further, the effect of time of the year on hospital visit occurrence was assessed. The Pearson's chi-square test for independence in the contingency table of number of patients in each of 12 months classified by the cluster was performed (Table 7). The Wilcoxon 2-sample test[36] was conducted to assess differences between winter and summer seasons.

RESULTS AND COMMENT

Emission sources

The emission inventory estimates (total annual emissions of SO_2, NO_x, and PM expressed as t [ton] or kg [kilogram]

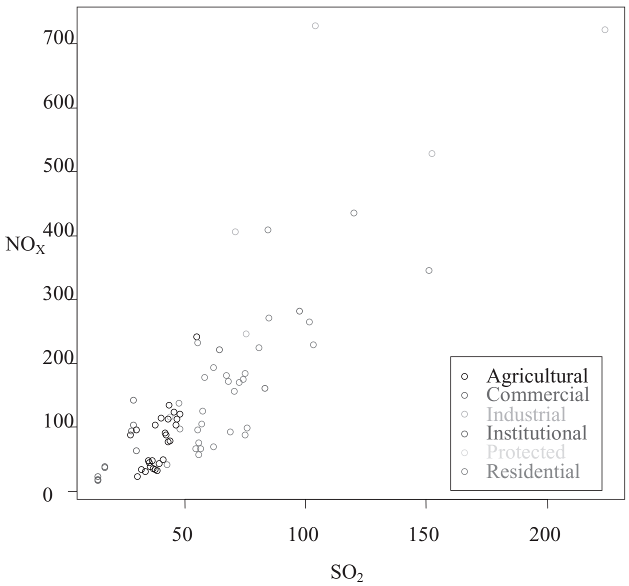

Fig. 4. Correlation between NO_x and SO_2 for land use types (unit: kg/day/grid) (color figure available online).

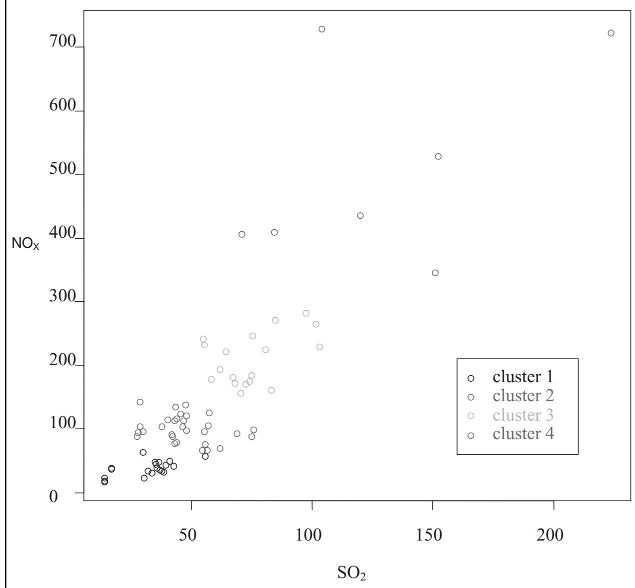

Fig. 5. Correlation between NO_x and SO_2 in emission clusters (unit: kg/day/grid) (color figure available online).

per day) showed that the main sources for outdoor air pollution are industries (37,147 kg/day), domestic fuel burning (8,455 kg/day), and vehicles (8,172 kg/day) (Supplementary Files 1, 2, 3).

Of total 7.6 t/day SO_2 emissions, industrial sources accounted for nearly 50% of total emission and the remaining 50% is attributed to other sources such as domestic fuel burning, vehicles, etc. (Supplementary File 1).

PM emissions accounted for 7.0 t/day. Forty-five percent of these emissions are attributed to domestic fuel burning. This source is followed by vehicles (20%), soil-road dust (18%), garbage burning (8%), and others (9%) (Supplementary File 2).

A total NO_x emission load of 19 t/day has been estimated within the city. The breakdown of emissions is as follows: industrial point sources (43%), vehicles (33%), domestic fuel burning (11%), industry area source (6%), diesel generator set (6%), and others (1%) (Supplementary File 3).

SO_2, NO_x, and PM emissions in grids are highly correlated as an artifact of the methodology, due to the fact that all 3 pollutants are emitted from most of the sources, and because each grid, even if classified into one land use type, usually has a mix of sources (Figure 4). From Figures 4 and 5, we can see that the values of SO_2 and NO_x are much higher in the industrial regions than in the institutional regions.

Emission clusters

The cluster analysis showed that the optimal number of clusters is either 2 or 4. For more exposure resolution, we chose to use 4 clusters. Table 4 provides emission levels for each variable of SO_2, PM, and NO_x within each of the 4 clusters. From Figure 5, we can see that the values of SO_2 and NO_x are much higher in the cluster 4 than in other clusters. Of the 154 grid cells, 78 that lie within the city limits were classified into 4 distinct pollution groups (ie, cluster) are seen in Figure 6.

Health assessment

Total Population in Each Emission Cluster

The spatial variation of number of inhabitants in each cluster is shown in Figure 7. Population is slightly lower in the very highly polluted regions (eg, a total of 232,224 people live in the regions within emission cluster 4). The highest number of inhabitants (748,457) lives in moderately highly polluted areas (ie, cluster 2; Figure 7).

Correlation Between Air Pollution and Respiratory Disease

In each level of emissions, the total populations, the total number of people that have visited hospital, and the total

Table 4.—Emission of SO_2, PM, and NO_x in Each Emission Cluster

Cluster number	SO_2 (kg/day/grid)	PM (kg/day/grid)	NO_x (kg/day/grid)
1	36.08	44.57	39.00
2	46.57	75.72	104.23
3	62.19	134.76	194.15
4	120.16	259.34	434.97

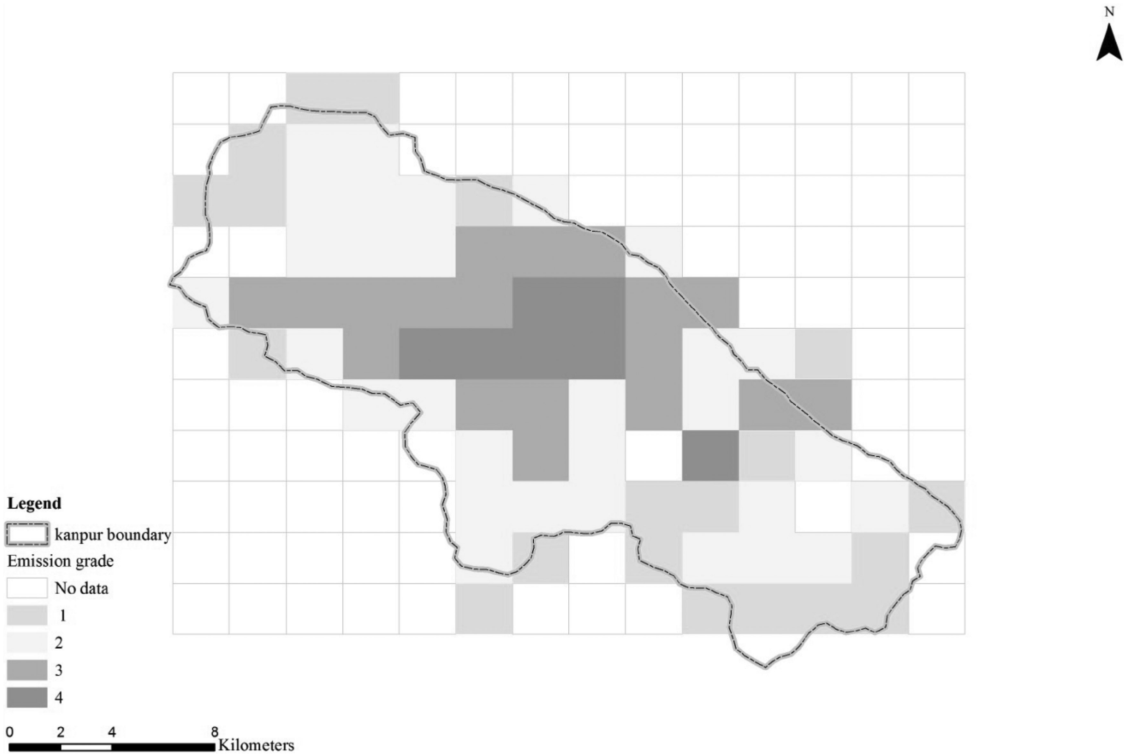

Fig. 6. Grid squares with indicated emission clusters (1 = less polluted, 2 = polluted, 3 = highly polluted, 4 = very highly polluted; refer to Table 4) (color figure available online).

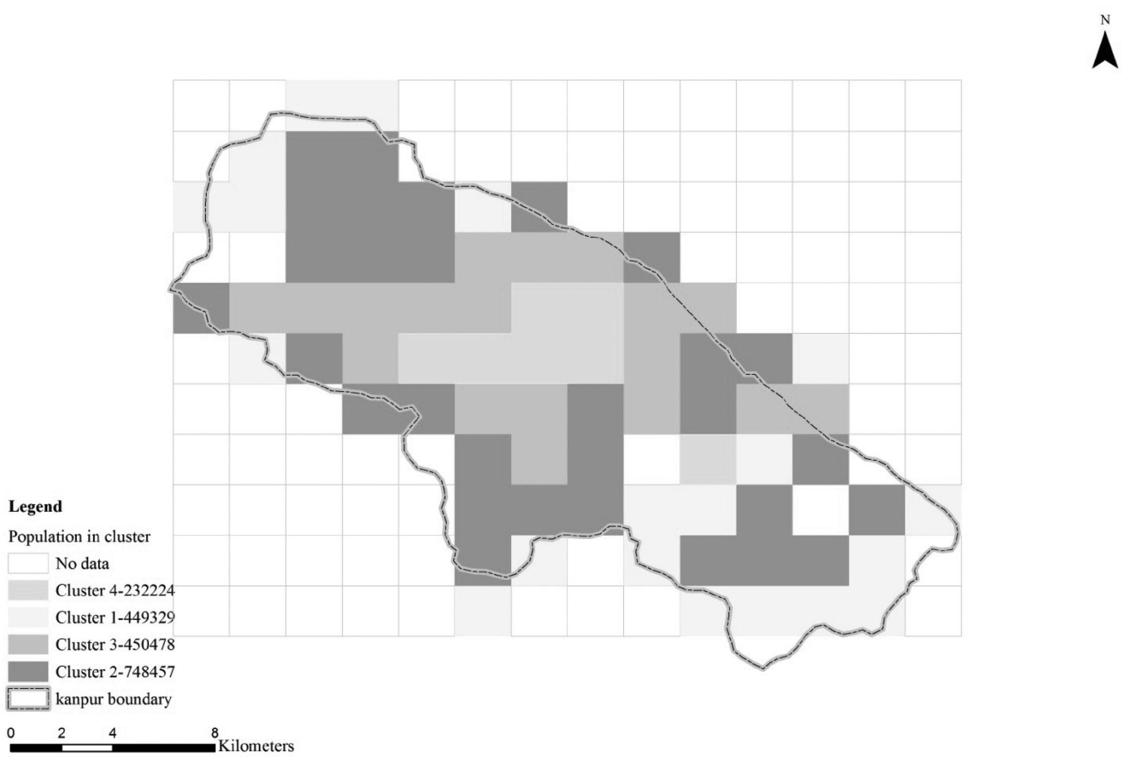

Fig. 7. Total number of inhabitants in each level of emissions (1 = less polluted, 2 = polluted, 3 = highly polluted, 4 = very highly polluted) (color figure available online).

Table 5.—Total Number of Inhabitants in Each Cluster Who Visited the LLR Hospital and the Total Number of the Control Group (see Data Analysis)

Patients	Cluster			
	1	2	3	4
Yes	466	1,274	1,553	781
No	448,863	747,183	448,925	231,443

number of people that have not visited the hospital are incorporated into this analysis (Table 5).

Pearson's chi-square test ($df = 3$, chi-square $= 835.52$, $p < .0001$) showed that the relative number of patients who visited the hospital per number of inhabitants in each cluster is much higher in the highly polluted regions (clusters 3 and 4; Figures 8, 9) than in the less polluted regions (cluster 1); the ratio in the polluted cluster 3 and 4 is higher than 0.003 (Figure 8).

To see to what extent the type of analysis affects the results, we have also used a logistic regression to model visits to the respiratory hospital in relation to the level of emissions. Cluster inhabitants not having visited the hospital represented the control group. The results were consistent: independent of the cluster variable type (eg, nominal, ordinal, or quantitative), the level of emission significantly affects the probability of visit to the hospital on any reasonable level ($p < .0001$; Table 6).

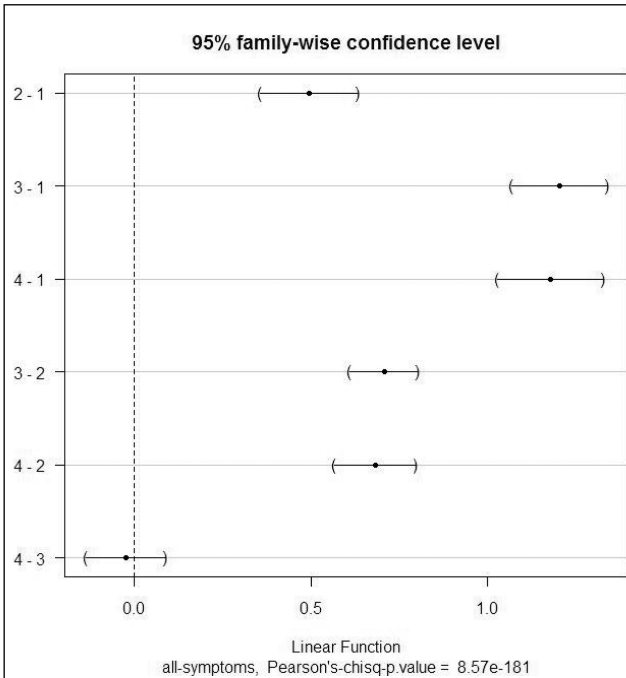

Fig. 9. Difference of effect on total symptom morbidity between each pairs of emission clusters (*y*-axis, each pair of emissions).

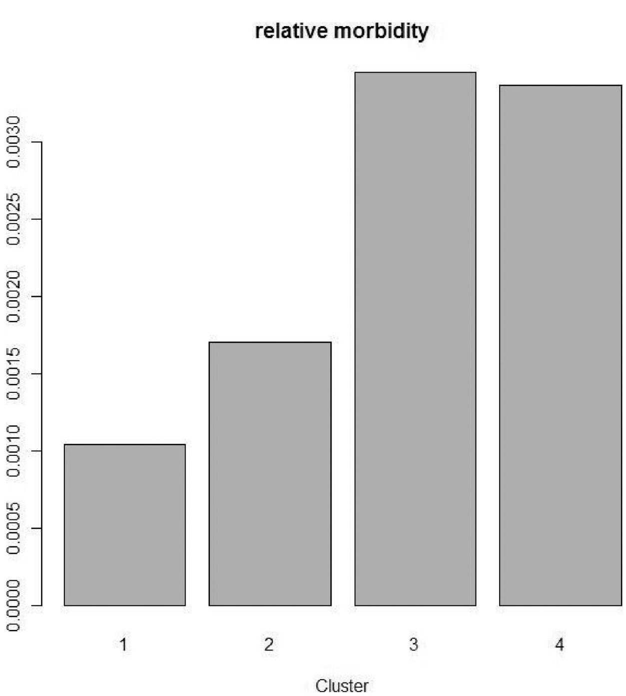

Fig. 8. The ratio of number of inhabitants who visited the LLR Hospital with respiratory symptoms to the total number of inhabitants ("relative morbidity") in the emission clusters (*y*-axis, relative morbidity).

The 95% simultaneous confidence interval (CI) ("family-wise") for the difference of effect on hospital visits between all the pairs of emission clusters indicated that this probability significantly differs between each pair of emission clusters (pairwise Pearson's chi-square test, $p < .001$), except between the highly polluted cluster 3 and very highly polluted cluster 4 (Figure 9; Pearson's chi-square test, $p = .59$). There is no discernible difference in effect between emission levels of clusters 3 and 4; it can be attributed to a saturation effect that beyond a certain pollution level, there is no significant increase in the probability to visit to the hospital.[37–39]

Table 7 provides further quantification of the above. It provides odds ratios (the value by which the relative risk of having respiratory disease is multiplied when we changed from one emission level to another), and 95% CIs of odds ratios. For instance, comparing emission clusters 1 (less polluted) and 3 (highly polluted), in the grids where the average pollution of SO_2 increases from 36.08 to 62.19 kg/day/grid, PM from 44.57 to 134.76 kg/day/grid, and NO_x from 39.00 to 194.15 kg/day/grid (Table 4), the relative risk of increasing respiratory diseases of the inhabitants is higher than 3.33 (95% CI: 2.91, 3.81)) (Table 7), which is a very high increase. This indicates that the respiratory disease morbidity is much higher in the highly polluted regions. This is consistent with findings that the levels of air pollution make a significant

Table 6.—Summary Results From Testing the Effect of Emission Classification Method (Quantitative, Ordinal, and Nominal Values for Exposure) on the Resulting Relationship With Hospital Visits Using Logistic Regression

Cluster (1, 2, 3, 4)		Estimate	*SE*	*z* value	*Pr* (>|*z*|)
Quantitative variable	c(1, 2, 3, 4)	0.42470	0.01599	26.57	<2e-16***
Ordinal variable	as.ordered(c(1, 2, 3, 4)).L	0.94894	0.04020	23.603	<2e-16***
	as.ordered(c(1, 2, 3, 4)).Q	−0.26049	0.03488	−7.469	8.07e-14***
	as.ordered(c(1, 2, 3, 4)).C	−0.21101	0.02857	−7.386	1.52e-13***
		df	Deviance	Residual deviance	*Pr* (>*Chi*)
Nominal variable	as.factor(c(1, 2, 3, 4))	3	830.48	0.00	2.2e-16***

Note. The null hypothesis is that all of the regression coefficients in the model are equal to zero. The z value is the Wald statistics; Pr (>|z|) is the probability level of 2-tailed test; df is degrees of freedom; Pr (>Chi) is p value that define the probability of observing a chi-square statistic at least as observed under null hypothesis. c = cluster; L = linear; Q = quadratic; C = cubic; *** = the probability level of 2-tailed test or chi-square test is significant.

contribution to the variation in daily hospital administration for respiratory disease.[1,2,3,39,40]

Similarly, by analyzing separately the individual groups of respiratory symptoms, we got the same result that emission levels 1 and 3 differ significantly for all symptoms, and there is no difference between emission levels 3 and 4 for any of the respiratory symptoms (Figure 10).

Analysis of Seasonal Differences

The aim of this analysis is to assess the effect of a certain time period (ie, months and seasons) on the distribution of relative number of inhabitants that visit hospital on each level of emissions, or an effect of each level of emissions on the distribution during the year. To avoid multiple testing problem and to clearly state what hypothesis we test, we chose month as time period (because this period could be appropriate, not too short or not too long, to capture seasonal effect on morbidity). We performed Pearson's chi-square test for independence in the hospital visits (Table 7) and obtained a p value of .29. Thus, there is not strong enough evidence to claim that the distribution of morbidity in the different emission levels depends on a certain month or that the distribution of morbidity during the year depends on a certain emission level. Table 8 presents total number of inhabitants who visited the LLR hospital in each month with residence in each of the emission clusters.

Another question we can ask is: Does the morbidity change during the year regardless of the emission level? Table 9 presents the number of patients for each month without considering the emission level. We divided 12 months into 2 samples. The first sample consists of the summer months (April to September) and the second sample consists of the remaining winter months. The Wilcoxon 2-sample test was conducted, with the result ($p = .015$) showing that there is a significant difference for respiratory disease morbidity between summer and winter months. It remains, however, not clear what seasonal factors affect the hospital visits (eg, extreme weather conditions in summer months, high pollution load in winter months), or if the hospital visits are mainly related to specific respiratory symptoms.

Table 7.—Odds Ratios Quantifying Increased Risk of Visiting the LLR Hospital Between the Emission Clusters

Contrasts between clusters	Estimate	95% CI	Odds ratio	95% CI
2 and 1	0.50	0.36, 0.63	1.64	1.43, 1.89
3 and 1	1.20	1.07, 1.34	3.33	2.91, 3.81
4 and 1	1.18	1.03, 1.33	3.25	2.80, 3.78
3 and 2	0.71	0.61, 0.80	2.03	1.84, 2.24
4 and 2	0.68	0.57, 0.80	1.98	1.76, 2.22
4 and 3	−0.02	−0.14, 0.09	0.98	0.87, 1.09

Note. CI = confidence interval.

Table 8.—Total Number of Inhabitants Who Visited the LLR Hospital in Each Month, With Residence in Each of the Emissions Cluster (for Patients With Identified Address in the Urban Area—78 Clusters)

	Cluster			
Month	1	2	3	4
1	30	95	129	57
2	44	99	128	71
3	45	111	161	66
4	44	106	139	55
5	45	98	121	73
6	36	107	131	71
7	32	106	133	75
8	47	126	127	69
9	29	119	131	77
10	28	71	109	62
11	48	113	117	52
12	19	70	88	38

Table 9.—Total Number of Patients in Each Month

	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
Total number of patients	555	684	738	680	747	784	772	837	768	581	732	462

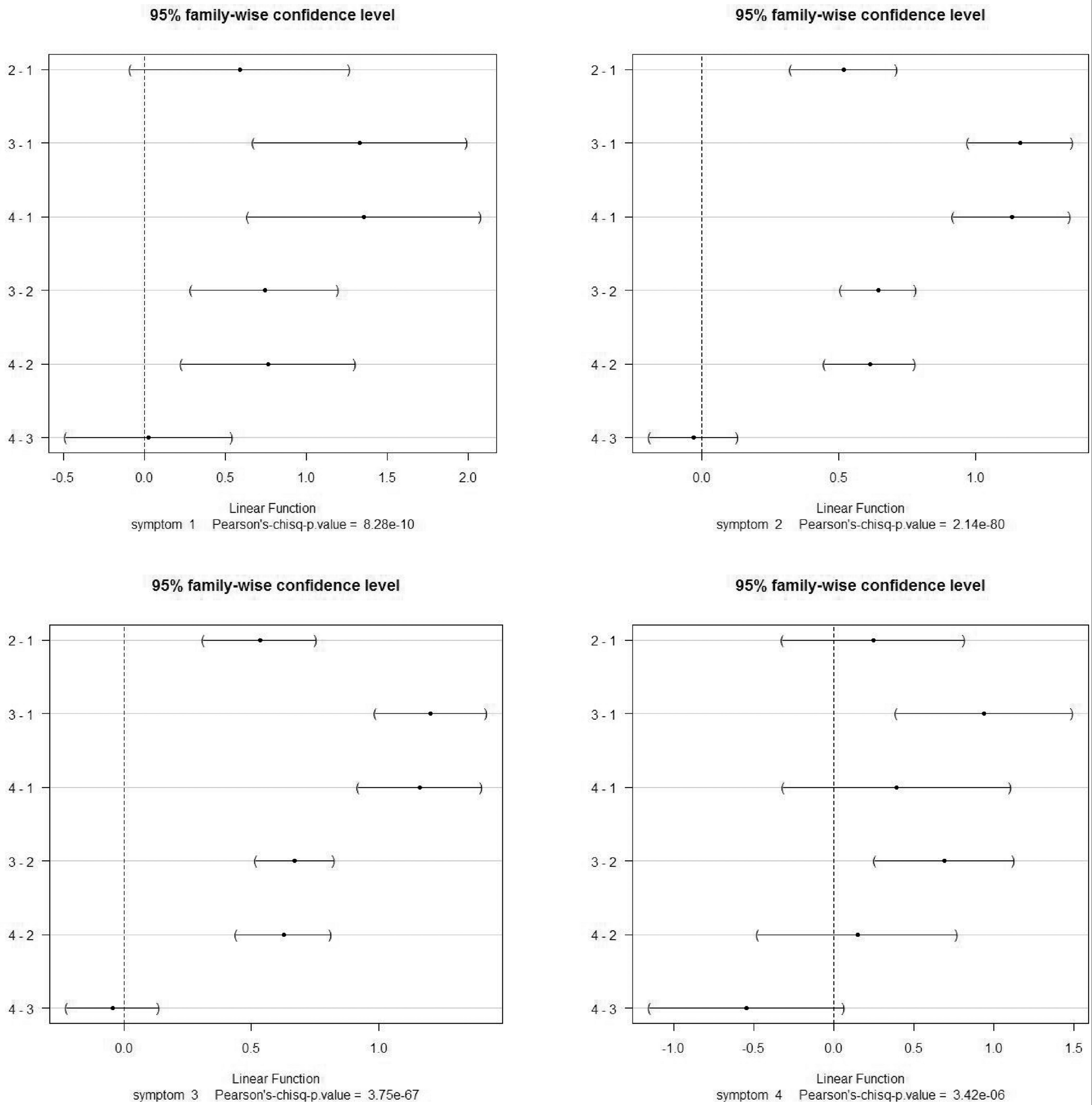

Fig. 10. Difference of effect on individual symptom morbidity between each pairs of emission clusters (for each classified symptom, see Table 3) (*y*-axis, each pair of emissions).

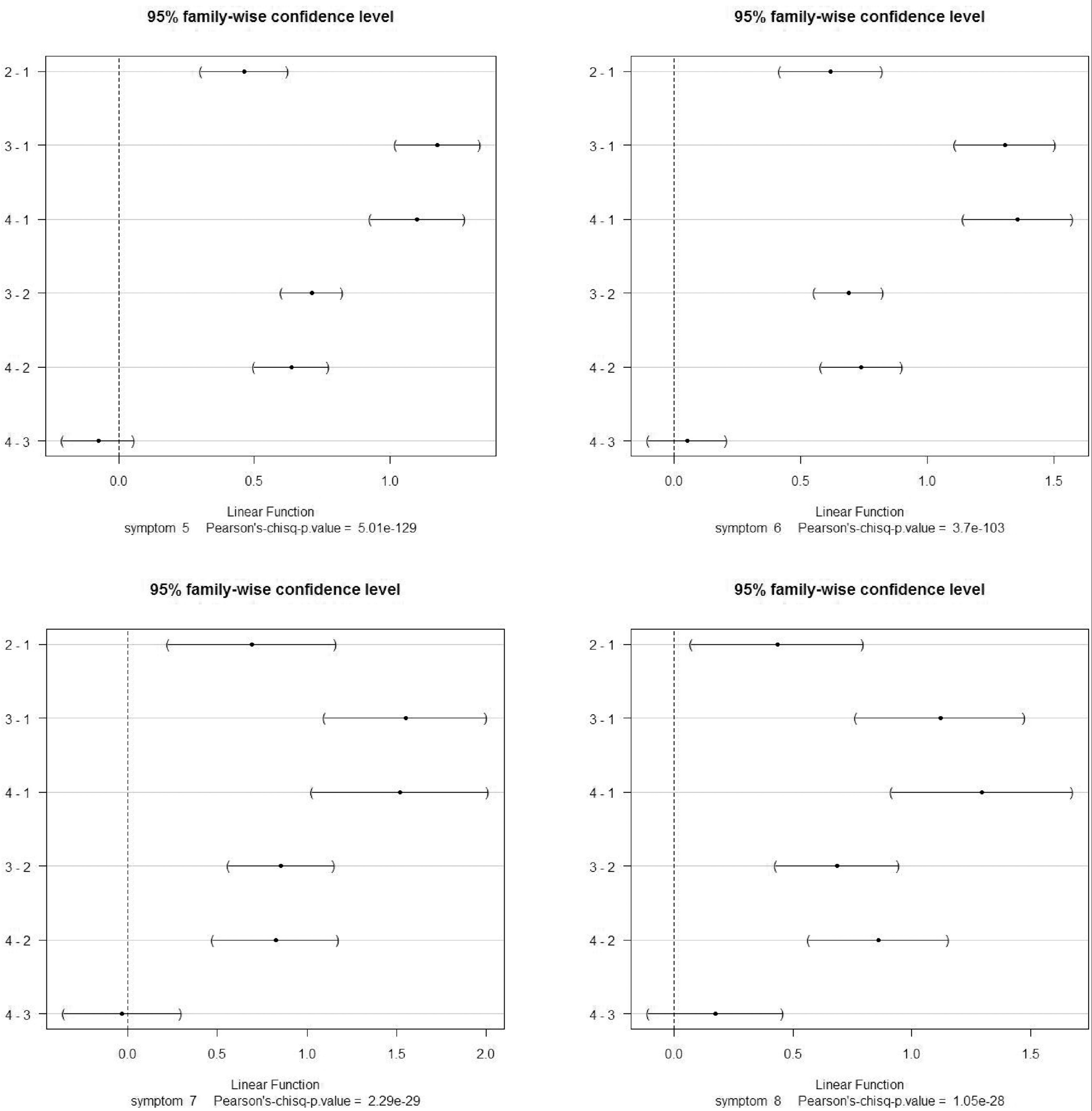

Fig. 10. (Continued)

Analysis of Occurrence for a Respiratory Disease Symptom

A logistic regression model with the occurrence of the symptom as the response and the variables cluster, sex, and age as the regressors showed that only the symptoms 4 (common cold), 5 (cough), and 6 (fever) are significantly impacted by the emission level. Common cold and cough are less frequent in the higher polluted regions, whereas fever occurs more likely in the higher polluted regions. However, symptom recording and classification is perhaps the part of the retrospective data collection that is most subjective and thus least rigorous, and this may affect the result.

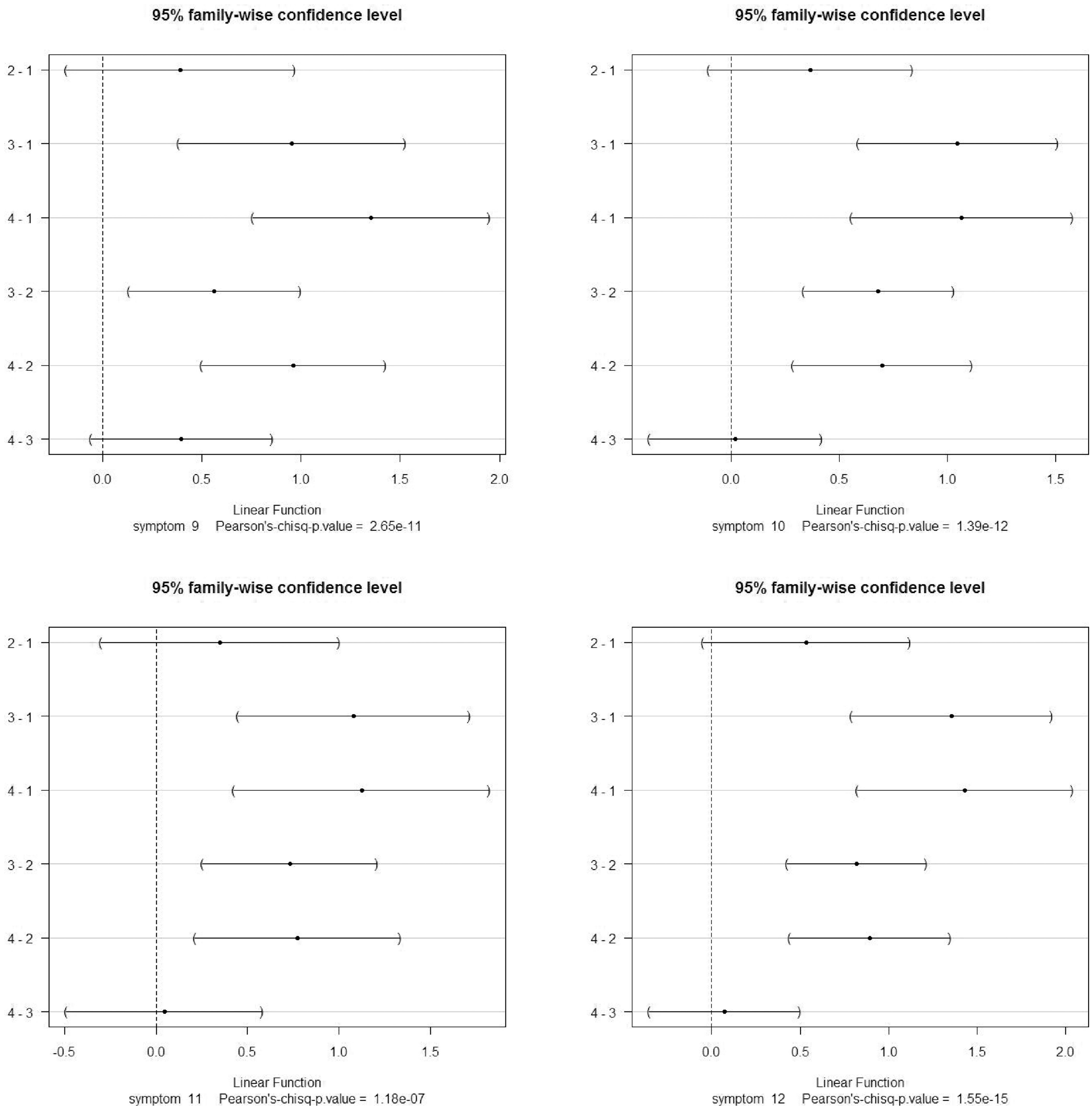

Fig. 10. (Continued)

Conclusions and recommendations

The main sources for outdoor air pollution in Kanpur are industries, domestic fuel burning, and vehicles. An emission inventory for the urban area was established, and local emissions in the 2 km $\times$ 2 km grid square representing the home address of Kanpur inhabitants were used as a surrogate measure of exposure to outdoor air pollution.

There is clear evidence that outdoor air pollution is associated with cardiopulmonary diseases. We did not investigate the effects on cardiovascular disease in this study, but we have shown that people living in the more polluted regions had a higher risk of hospital visits related to respiratory diseases than those living in the less polluted areas. This phenomenon is in accordance with other study results.[38,40–45]

Our study has several limitations. The exposure is represented by emission category (cluster) for each individual home. This has both advantages (being a more robust method), and disadvantages; it forces the analysis into a cross-sectional mode, considering a cumulative or long-term effect of pollution on respiratory health.[46] The retrospective data collection, especially to record respiratory symptoms, poses also some difficulties.

Our study could not consider time lags between air pollution and the occurrence of respiratory disease morbidity[47,48] due to a lack of temporal data on air pollution. Despite these shortcomings, we did show a clear relationship between pollution and respiratory disease. Clearly, more information is available in the data than we have used in this analysis, but we feel that our result is robust, and should be sufficient to trigger further work. In the future, this analysis can be supplemented by adding daily variability in pollution concentrations, allowing an analysis of short-term effects.

All the findings from this study suggest that long-term, systematic, prospective epidemiological studies on exposure to air pollution and its respiratory health effects are needed.[49]

This paper indicates that the polluted air was a clear threat to human health in Kanpur in 2006. This situation is alarming. Mitigation will require a combination of technical measures on all sources (this study was conducted before the massive conversion of the public transport to compressed natural gas) and urban planning. Also, it is necessary to establish adequate monitoring systems both for air quality and for human health, to be able to assess the trends, and the effectiveness of measures taken.

****

This study has been carried out under the project "Environmental Health Assessment: Respiratory Disease in relation to Air Pollution in Kanpur, Uttar Pradesh," which was an 18-month cooperation project between the Indian Institute of Technology Kanpur (IITK), Ganesh Shankar Vidyarthi Memorial (GSVM) Medical College in Kanpur, State Pollution Control Board (SPCB) in Kanpur, and Norwegian Institute for Air Research (NILU), funded under Royal Norwegian Embassy New Delhi grant (IND 3025 05/051). In addition to the partners funded by this grant, the Central Pollution Control Board, Agra, participated in the project's intercomparison exercises and preparation of standard operating procedures (SOPs).

The authors would like to thank a number of colleagues for their invaluable contribution. The staff at the laboratory at CPCB and the students and staff at the IITK have been instrumental in all logistics and technical preparation, quality control, as well as monitoring activities. The authors would like to thank the station managers who diligently supported their efforts. The authors thank especially Mike J. Kobernus and Sonja Grossberndt at NILU for help with the language.

For comments and further information, address correspondence to Alena Bartonova, Norwegian Institute for Air Research (NILU), Instituttveien 18, Postboks 100, 2027 Kjeller, Norway.

E-mail: alena.bartonova@nilu.no

****

References

1. Pope CA III, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. Adults. *Am J Respir Crit Care Med*. 1995;151:669–674.
2. Xu X, Li B, Huang H. Air pollution and unscheduled hospital outpatient and emergency room visits. *Environ Health Perspect*. 1995;103:286–289.
3. Brook RD, Rajagopalan S, Pope CA III, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*. 2010;121:2331–2378.
4. Chen B, Kan H. Air pollution and population health: a global challenge. *Environ Health Prevent Med*. 2008;13:94–101.
5. Sivertsen B. Global ambient air pollution concentrations, trends. In: World Health Organization, ed. *Air Quality Guidelines, Global Update 2005*. Copenhagen, Denmark: WHO Regional Office for Europe; 2006:31–59.
6. Medina S, Le Tertre A, Saklad M. Apheis Collaborative Network. The Apheis project: Air Pollution and Health–A European Information System. *Air Qual Atmos Health*. 2009;2:185–198.
7. Stieb DM, Judek S, Burnett RT. Meta-analysis of time-series studies of air pollution and mortality: effects of gases and particles and the influence of cause of death, age, and season. *J Air Waste Manage Assoc*. 2002;52:470–484.
8. Katsouyanni K, Touloumi G, Samoli E, et al. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology*. 2001;12:521–531.
9. Arbex MA, de Souza Conceic GM, Cendon SP, et al. Urban air pollution and chronic obstructive pulmonary disease-related emergency department visits. *J Epidemiol Community Health*. 2009;63: 777–783.
10. Tramuto F, Cusimano R, Cerame G, et al. Urban air pollution and emergency room admissions for respiratory symptoms: a case-crossover study in Palermo, Italy. *Environ Health*. 2011;10:1–11.
11. Miller KA, Siscovick DS, Sheppard L, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med*. 2007;356:447–458.
12. Andersen ZJ, Hvidberg M, Jensen SS, et al. Chronic obstructive pulmonary disease and long-term exposure to traffic-related air pollution: a cohort study. *Am J Respir Crit Care Med*. 2011;183: 455–461.
13. Wong CM, Atkinson RW, Anderson HR, et al. A tale of two cities: effects of air pollution on hospital admissions in Hong Kong and London compared. *Environ Health Perspect*. 2002;110:67–77.
14. Hwang JS, Chan CC. Effects of air pollution on daily clinic visits for lower respiratory tract illness. *Am J Epidemiol*. 2002;155:1–10.
15. Health Effects Institute (HEI). *Health Effects of Outdoor Air Pollution in Developing Countries of Asia: A Comprehensive Review*. Boston: Health Effects Institute; 2010. Report number 18.
16. Wong C-M, Vichit-Vadakan N, Kan H, Qian Z, the PAPA Project Teams. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. *Environ Health Perspect*. 2008;16:1195–1202.
17. Chung KF, Zhang J, Zhong N. Outdoor air pollution and respiratory health in Asia. *Respirology*. 2011;16:1023–1026.
18. Sharma M, Maloo S. Assessment and characterization of ambient air PM_{10} and $PM_{2.5}$ in the City of Kanpur, India. *Atmos. Environ.* 2005;39:6015–6026.
19. Sharma M, Narendra KV, Katiyar SK, Sharma R, Shukla BP, Sengupta B. Effects of particulate air pollution on the respiratory health of subjects who live in three areas in Kanpur, India. *Arch Environ Health*. 2004;59:348–358.
20. Kumar R, Sharma M, Srivastva A, Thakur JS, Jindal SK, Parwana HK. Association of outdoor air pollution with chronic respiratory morbidity in an industrial town in northern India. *Arch Environ Health*. 2004;59:471–477.
21. Nath KJ. Environmental health monitoring and management action plan. In: WHO Regional Office for South-East Asia, ed. *Environmental Epidemiology in the South-East Asia Region—Report of a Regional Training*. New Delhi, India: World Health Organization; 2000:7–8.
22. Bartonova A, Sharma M, eds. *Indoor and Ambient Air Exposure of PAHs and Fine Particulate to Women and Children: Health Impacts in terms of Morbidity*. Kjeller, Norway: Norwegian Institute for Air Research; 2005.
23. Kumar R, Nagar JK, Kumar H, et al. Association of Indoor and outdoor air pollutant level with respiratory problems among children in an industrial area of Delhi, India. *Arch Environ Occup Health*. 2007;62:75–80.

24. Rosati JA, Yoneda KY, Yasmeen S, Wood S, Eldridge MW. Respiratory health and indoor air pollution at high elevation. *Arch Environ Occup Health*. 2005;60:96–105.
25. Singh RP. Effect of wastewater disposal and extent of industrial pollution in and around Kanpur, Uttar Pradesh, India. *Bull Eng Geol Environ*. 2001;60:31–35.
26. Sharma M. *Air Quality Assessment, Emissions Inventory and Source Apportionment Studies for Kanpur City*. New Delhi, India: Central Pollution Control Board; 2010.
27. Anonymous. R version 2.7.1 for Windows. The Comprehensive R Archive Network Web site. Available at: http://cran.r-project.org/bin/windows/base/old/2.7.1/.Updated June 23, 2008. Accessed October 7, 2004
28. Anonymous. R: a language and environment for statistical computing. Mendeley Web site. Available at: http://www.mendeley.com/research/r-a-language-and-environment-for-statistical-computing-2/. Accessed October 7, 2008.
29. Oksanen J, Kindt R, Legendre P, O'Hara B, Henry M, Stevens H. The vegan package. University of Murcia Web site. Available at: http://ocw.um.es/ciencias/geobotanica/otros-recursos-1/documentos/vegan.pdf. Accessed February 16, 2008.
30. Kirk RE, ed. *Experimental Design: Procedures for the Behavioral Sciences.* 3rd ed. Pacific Grove, CA: Brooks/Cole Publishing Company; 1995.
31. Oksanen J. Cluster analysis: tutorial with R. University of Oulu Web site. Available at: http://cc.oulu.fi/∼jarioksa/opetus/metodi/sessio3.pdf. Accessed March 6, 2012.
32. Greenwood PE, Nikulin MS, ed. *A Guide to Chi-Squared Testing*. New York: Wiley; 1996.
33. Manning C. Logistic regression (with R). The Stanford University Web site. Available at: http://nlp.stanford.edu/∼manning/courses/ling289/logistic.pdf. Accessed March 6, 2012.
34. Lemeshow S, Hosmer DW, ed. *Applied Logistic Regression.* 2nd ed. New York: Wiley; 1989.
35. Kleinbaum DG, ed. *Logistic Regression: A Self-Learning Text.* 2nd ed. New York: Springer; 1994.
36. Anonymous. Wilcoxon rank sum and signed rank tests. Seminar for Statistics Web site. Available at: http://stat.ethz.ch/R-manual/R-patched/library/stats/html/wilcox.test.html. Accessed March 6, 2012.
37. Petroeschevsky A, Simpson RW, Thalib L, Rutherford S. Association between outdoor air pollution and hospital administrations in Brisbane, Australia. *Arch Environ Health*. 2001;56:37–52.
38. Patel S, Nellore MRR, Sadhu HG, Kulkarni PK, Patel BD, Parikh DJ. Effects of Industrial pollution on respiratory morbidity among female residents of India. *Arch Environ Occup Health*. 2008;63:87–92.
39. Sunil K, Chhabra SK, Pragti Chhabra P, Gupta RK. Ambient air pollution and chronic respiratory morbidity in Delhi. *Arch Environ Health*. 2010;56:58–64.
40. Segala C. Health effects of urban outdoor air pollution in children. *Pediatr Pulmonol (Suppl)*. 1999;27:6–8.
41. Sagar A, Bhattacharya M, Joon V. A comparative study of air pollution-related morbidity among exposed population of Delhi. *Indian J Community Med*. 2007;32:10–12.
42. Nawrot TS, Alfaro-Moreno E, Nemery B. Update in occupational and environmental respiratory disease 2007. *Am J Respir Crit Care Med*. 2008;177:696–700.
43. Mansourian M, Javanmard SH, Poursafa P, Kelishadi R. Air pollution and hospitalization for respiratory diseases among children in Isfahan, Iran. *Ghana Med J*. 2010;44:138–143.
44. Chen H, Goldberg MS. The effects of outdoor air pollution on chronic illnesses. *McGill J Med*. 2009;12:58–64.
45. Forbes LJL, Patel MD, Alicja R, Rudnicka AR, et al. Chronic exposure to outdoor air pollution and diagnosed cardiovascular disease: meta-analyis of three large cross-sectional surveys. Environ Health. 2009;8:30.
46. Brauer M. How much, how long, what, and where–air pollution exposure assessment for epidemiologic studies of respiratory disease. *Proc Am Thorac Soc*. 2010;7:111–115.
47. Kalantzi EG, Makris D, Duquenne MN, Kaklamani S, Stapountzis H, Gourgoulianis KI. Air pollutants and morbidity of cardiopulmonary diseases in a semi-urban Greek peninsula. *Atmos Environ*. 2011;45:7121–7126.
48. Ignotti E, Hacon SDS, Junger WL, et al. Air pollution and hospital admissions for respiratory diseases in the subequatorial Amazon: a time series approach. *Cadernos Saúde Pública*. 2010;26:747–761.
49. Öry FG, Rahman FU, Katagade V, Shukla A, Burdorf A. Respiratory disorders, skin complaints, and low-back trouble among tannery workers in Kanpur, India. *Am Ind Hyg Assoc J.* 1997;58:740–746.