Alternative models in the analysis of the association between air pollution and respiratory health

Giuseppe Migliaretti, Dario Gregori, Paola Berchialla, Franco Cavallo

Department of Public Health and Microbiology, University of Turin, Italy

Corresponding Author:
Giuseppe Migliaretti
University of Turin, Dept of Public Health, Statistical Unit
Via Santena 5bis, Turin, Italy
E-mail: giuseppe.migliaretti@unito.it

Summary

The aim of this study was to investigate the association between urban air pollution and emergency hospital admissions for respiratory causes in the population resident in Turin in the years 1998-1999, using a case-control design. *Cases* were defined as those patients resident in Turin and admitted for chronic bronchitis, emphysema, other chronic obstructive pulmonary diseases, or asthma [n_1 = 3.997]; for each case, up to 30 *controls*, matched for gender, age and admission year, were randomly selected among patients resident in Turin and admitted for causes other than respiratory or heart diseases [n_2 = 124.005]. Consistently with the literature, the levels (in μg/m^3) of total suspended particulate (TSP) and nitrogen dioxide (NO2) were taken as the principal indicators of urban air pollution. In our study, we considered each patient's exposure to each pollutant on the admission day and in the previous two weeks. To analyze the effects of these pollutants on hospital admissions for respiratory health problems, an exploratory and an associative analysis, based on projection pursuit regression (PPR) and conditional logistic regression models respectively, were performed. For each pollutant, the effect was expressed as the percent increase in risk (and relative 95% CI) per 10 μg/m^3 increase in pollutant concentration. The PPR model results showed, for TSP and NO2, a principal effect occurring between 7 and 10 and between 8 and 12 days prior to admission, respectively. On the basis of these results, the mean level computed in the period 7-10 days for TSP, and in the period 8-12 days for NO2, was taken, for each patient, as the indicator of exposure to the respective pollutant. The logistic regression model results showed a mean increase of 3.1% (95% CI 2.3-3.9%) and of 5.1% (95% CI 3.5-6.8%) for TSP and NO2 respectively. Despite some limitations, the case-control approach used in our study appears to be useful for estimating the respiratory illness admission risk associated with air pollution exposure and the results confirm that the effect occurs in the days prior to the admission, underlining the need for further research in this area.

KEY WORDS: *air pollution, respiratory health, case-control, logistic regression, projection pursuit regression (PPR).*

Introduction

In the literature there is growing evidence of the effects of air pollution on the incidence of cancer, respiratory and cardiovascular diseases, and on the rise of general morbidity (1-8).

Defining the potential effects of environmental pollution on health remains a thorny issue as regards both the attribution of exposure and the identification of many confounding factors, which are often difficult to control. However, hospital admissions for respiratory causes are a usual outcome of the short-term effects of air pollution. Close correlations between increases in air pollution and numbers of hospital admissions for respiratory causes have been shown by studies conducted in the US (9-15) and also by several European studies (16-24).

The statistical methods used in most of these studies were based on time-series data to control for associations between long-term seasonal variation effects and/or cyclical components and changes in exposure and frequency of health events (23, 24). Very few studies have used a case-control design to evaluate the relationship between hospital admissions and urban air pollution. Ciccone et al., who investigated the population of Ravenna, Italy (25), Panella et al. who considered the population of Novara (Italy) (26), and Migliaretti et al., who considered the population of

Turin (27), all used a case-control design to analyze the relationship between numbers of hospital admissions and levels of exposure to urban pollutants. However, all the studies investigating the risk of admission for respiratory illness associated with air pollution exposure took the mean exposure level estimated in the days immediately prior to admission (frequently the day of admission itself and the three days before admission) as an indicator of total exposure; few data are available about the effects of exposure over a longer period before admission. A recent Italian study (28) showed different effects of pollutants in the 15 days before admission; the PM10 effect decreased, but a more irregular pattern was found for NO2 and SO2.

In the light of the results published in the literature, further investigations could be useful, using alternative approaches both for estimating hospital admission risk and for analyzing the effect, on respiratory admissions, of single pollutants measured in the days prior to the admission day.

Objectives

Using a case-control design based on routinely collected data, our study investigated the association between urban air pollution and emergency hospital admissions for respiratory causes in the population resident in Turin in the period from January 1998 to December 1999, analyzing the effect of the exposure to total suspended particulate (TSP) and nitrogen dioxide (NO2) in the two weeks prior to admission.

Materials and methods

Admission data

Admission data were extracted from the regional archive of hospital discharge records, routinely kept by the epidemiology service of the Piedmont Region. The study was based on emergency admissions of patients living, as residents, in the city of Turin in the period from January 1998 to December 1999. Patient information included demographic data, place of residence and details of the admission. Based on the patient's principal diagnosis, coded according to the International Classification of Diseases, 9th Revision (ICD-9) (29), *cases* were defined as those patients resident in Turin and admitted for chronic bronchitis, emphysema, or other chronic obstructive pulmonary diseases (ICD-9 codes 490--492 and 494--496), or asthma (ICD-9 code 493) [$n_1 = 3.997$]; for each case, up to 30 *controls*, matched for gender, age and admission year, were randomly selected among patients, also resident in Turin, admitted for causes other than respiratory diseases (ICD-9 codes 460–487 and 500–519) or heart diseases (ICD-9 codes 390–405 410–429) [$n_2 = 124.005$].

Air-pollution and meteorological data

Consistently with the literature (17, 22, 24, 27), we considered, as indicators of urban air pollution, the average daily levels of TSP and NO2 (in both cases measured in $\mu g/m^3$).

Air-pollution data were provided by the Air Quality Control Division of the Regional Agency for Environmental Protection in Piedmont (ARPA). During the period from January 1997 to December 1999, three out of 10 monitoring stations located in different quadrants of Turin furnished reliable data for the parameters (TSP and NO2 levels) considered in this study.

As suggested by other studies on the short-term effects of air pollution (17, 24), the following indicators were taken as the ones best able to represent health risk: 24-hour maximum for NO2, 24-hour total for TSP. In our study, we considered each patient's exposure to each pollutant on the day of admission and in the two weeks before admission.

The completeness criteria for data recorded at the monitoring stations were similar to those used in other studies on the effects of pollution on health (17): stations lacking over 25% of data for the entire study period were excluded, and the daily data were defined as lacking when less than 75% of the daily hourly data were recorded. When data were not available in one or more stations on certain days, these data were substituted with the weighted mean of the values recorded from all the other monitoring stations, when available; i.e.: $X_{estimate} = X_d \times (X_{kj} / X_k)$, where: X_d = daily mean of the other monitoring sites,

X_{kj} = k yearly mean in the j monitoring site, X_k = mean for the k year of all the monitoring sites.

When the information was not available for all the monitoring stations, the days with missing values were replaced with the average of the value of the pollutant recorded during the previous and the subsequent day, if these data were not missing as well. In all other cases, data were reported as missing.

The Air Quality Division of ARPA also supplied meteorological data, drawn from their daily records. Mean daily temperature and humidity were the indicators selected on the basis of previous epidemiological evidence (17, 24); other meteorological data used in the study were wind velocity and solar radiation.

Statistical analysis

To estimate the effect of each pollutant on the admissions for respiratory health, considering both the admission day and the two weeks before the admission, an exploratory analysis was performed using a set of projection pursuit regression (PPR) models (30). PPR models were developed to map high-dimensional data into a lower dimensional space, typically a line or a plane, in an attempt establish their intrinsic structure. The PPR procedure consists of approximating the mean regression function among a set of covariates and a response variable using a sum of unknown ridge functions. In particular, given a data set **X** of k-covariates and a response *Y*, the PPR model is defined as:

$$Y = \beta_0 + \sum_{j=1}^{M} \beta_j f_j(\alpha_j X) + \varepsilon ,$$

where $\beta = (\beta_1, ..., \beta_M)$ are the regression coefficients, $\alpha = (\alpha_1, ..., \alpha_M)$ the k direction projections, and $f = (f_1, ..., f_M)$ the ridge functions.

Using the PPR models, non-linear effects on the probability of hospitalization of the principal components of the observed levels for each pollutant, adjusted by the other known factors, were estimated. Since the results of PPR models are dependent on the number of projection directions chosen, a set of PPR models with N=1, 2, 3 projection directions was implemented. The fitting procedure consists of building up a model with a maximum number of terms M_{max} and pruning back to N by dropping the least effective term and refitting. Thus, for each of the three PPRs, a set of models with M_{max} ranging from 8 to 18 was tried. The final model was then selected using the Akaike information criterion (AIC).

In the field of pollution series, PPR can be used as an exploratory tool to detect non-linear or linear trends in the series. Projection pursuit regression can be viewed as a general additive model applied to the projected variables, furnishing information about the projection directions (the covariates' weights), the ridge functions and finally the projection weights. Once such weights are plotted against the observed levels of pollution, a robust graph of the admission risk-pollution relationship can be obtained. The usefulness of this technique is that it is able to cope with very high dimensions, without the need to worry about collinearity of the selected covariates, like those coming from a sequential sampling of air concentrations.

Analyzing the results available from PPR models, for each pollutant, the period of increased effect on the admission event was defined as the period in which the admission risk-pollution relationship was higher than the mean value computed for the two weeks prior to the admission. In the analysis, for each pollutant, the mean level measured in the period of increased effect was taken as the indicator of the patient's exposure.

The exposure effects were analyzed using the conditional logistic regression model for matched case-control data and the results shown as percent increase in risk (IR%) and relative 95% Confidence Interval (95% CI) per 10 $\mu g/m^3$ increase in pollutant concentration. The analysis was performed separately for the different age groups of patients: children (<15 years), adults (15-64 years) and the elderly (>64 years).

As mentioned above and confirmed by the literature (31), the airborne pollutants were influenced by variations in weather (seasonality). Figure 1 shows the pollutant levels measured in the two-year period considered in this study. In the analysis, the month of admission, recoded as *Winter* (October-March) or *Summer* (April-September), was considered as a seasonality indicator (17).

Other adjustment factors considered in the analysis were mean daily temperature and mean daily humidity (wind velocity was not considered as the preliminary analysis showed it not to be influential); the different probability of admission according to the day of admission was controlled for the day of the week on which the patient was admitted. Since information about the subject's tobacco use and social status were unavailable from the admission data, educational levels were used to compensate partially for this gap. Statistical analysis was carried out using R software (32).

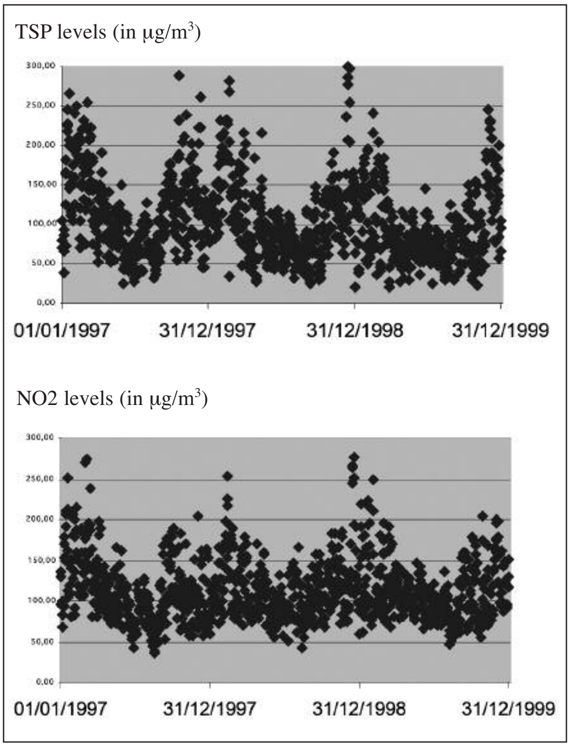


Figure 1. Total suspended particulate (TSP) and nitrogen dioxide (NO2) levels by day of recording in the period 1997-1999.

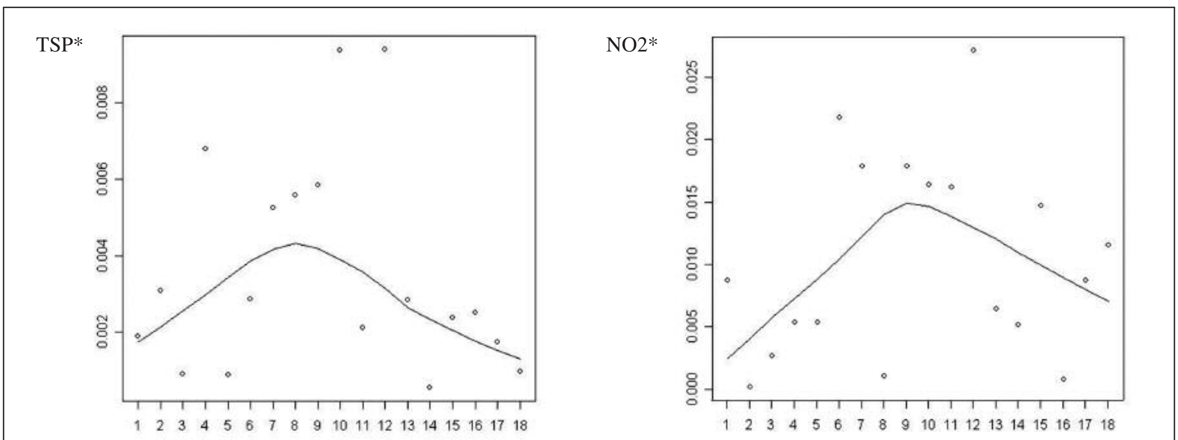
Results

Table 1 shows the performance of the PPR models. The model with 18 terms fitted sequentially and pruned back to two direction projections showed the best goodness of fit. It also showed the best AIC value. The results of this PPR model (Figure 2) showed, for TSP and NO2, a bigger effect between 7 and 10 days and between 8 and 12 days before the admission, respectively, with a tendency to a further increase (only for NO2) after 15 days. According with these results, for each patient, the indicator of exposure to TSP and to NO2 was taken as the mean level com-

Table 1. Projection pursuit regression models.

	Number of direction projections N		
M_{max} terms	1	2	3
8	0.48	0.33	0.36
9	0.51	0.27	0.45
10	0.48	0.30	0.36
11	0.45	0.24	0.30
12	0.54	0.36	0.42
13	0.51	0.30	0.45
14	0.51	0.27	0.39
15	0.60	0.27	0.33
16	0.60	0.24	0.38
17	0.48	0.22	0.36
18	0.48	0.21	0.30
AIC	5782.13	5607.76	5842.8

The procedure fits M_{max} terms sequentially, then prunes back to N the number of direction projections chosen by dropping the least effective term and re-fitting the model at each stage.



* In the x-axis the number of days before the admission and in the y-axis the weight of the pollutant.

Figure 2. Effects of total suspended particulate (TSP, in µg/m^3) and nitrogen dioxide (NO2, in µg/m^3) levels on hospital admissions for respiratory causes in the two weeks before admission.

Table 2. Percentage of increased risk (and relative 95% confidence interval) of hospital admission for respiratory causes per 10 μg/m^3 increase in exposure to the two pollutants considered.

Age category	TSP			NO2		
		95%CI			95%CI	
	IR%	Lower	Upper	IR%	Lower	Upper
All	3.10	2.30	3.90	5.10	3.50	6.80
≤ 15 years	0.50	-0.010	2.30	5.30	1.80	8.80
15-64 years	1.80	0.100	3.50	3.50	0.00	7.00
> 64 years	4.20	3.100	5.20	5.40	3.30	7.60

Abbreviations: TSP=total suspended particles; NO2=nitrogen dioxide.

puted in the period 7-10 days and 8-12 days prior to admission, respectively.

The logistic regression model revealed a significant association between a general increase in the risk of emergency admission and an increase in exposure (Table 2); in particular, the results showed a mean increase of 3.1% (95% CI 2.3-3.9%) and 5.1% (95% CI 3.5-6.8%) for TSP and NO2, respectively. The risk increased with the patient's age for TSP, while the effect of NO2 exposure was more evident among the young and elderly patients.

Discussion and conclusions

Our study on the population resident in Turin admitted to hospital from 1998 to 1999 showed an important effect of the exposure to pollutants, also many days before the admission, in particular for NO2 and TSP, and revealed a positive association between emergency admissions for respiratory causes and exposure to each pollutant considered. The association was particularly evident among the elderly patients for both pollutants, and also among the young for NO2.

Studies on the effects of atmospheric pollution on health must contend with some thorny problems, primarily the heterogeneity of pollutants, and the difficulty in measuring them accurately. Several groups of common air pollutants in urban air have been the object of studies conducted by the International Agency for Cancer Research (IARC). However, not even the monitoring of these pollutants by an air quality control network allows direct measurement of all of their components. Not even TSP, for example, can accurately represent the particles deriving from combustion in thermal plants and diesel exhausts, where the weighted mass consists mostly of coarse particles, often of biological origin, or originating from the erosion of crystal material. In this respect, no air quality parameter currently measured can be considered to cover, completely, the complex and variable "mixture" that is summed up by the term "urban air pollution".

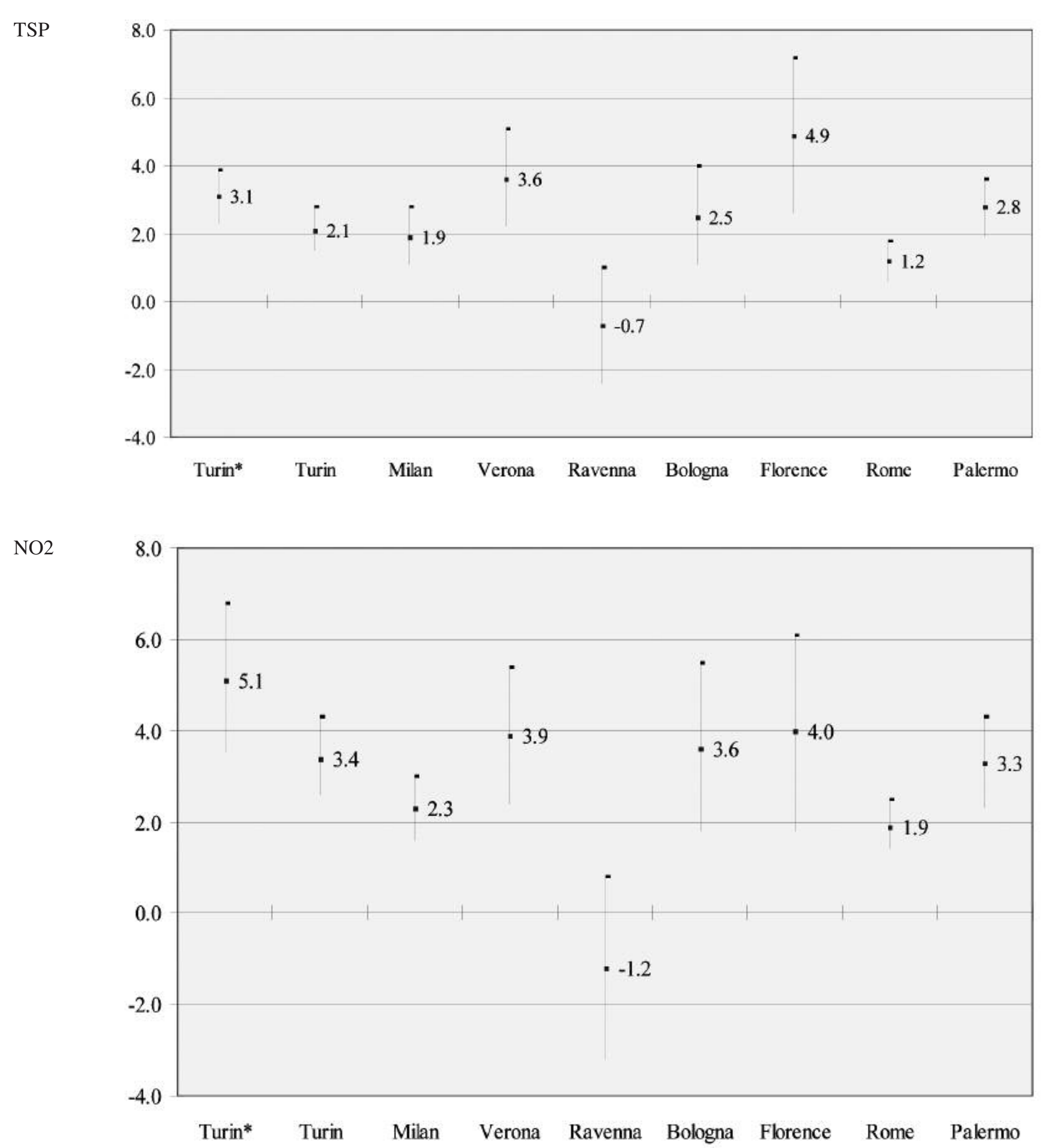
In our study, we hypothesized a general homogeneity of pollution levels across Turin during the course of the day, and thus attributed each subject with the same level of exposure to the pollutants. This choice was justified by the city's geographical characteristics; compared with other cities, Turin presents, across its whole area, a more homogeneous profile of meteorological and air pollution levels throughout the day (24, 28); for this reason the average daily measure of each pollutant can be considered a good indicator of the exposure for each patient in the study.

An important confounding factor not considered in the study was indoor exposure, but from this perspective, too, the city of Turin presents are marked heterogeneity, since most residences are centrally heated and most use gas cooking facilities (33); as a result, indoor exposure levels should not affect the results.

However, in interpreting the results of this study, consideration should be given to errors in the measurement of the pollutant predictors considered, as well as the possible effect of active and passive smoking exposure (the regional archive of hospital discharge forms did not permit an exact assessment of each subject's smoking status, and the educational level considered in the analysis can compensate only partially for this gap), and the risk of diasease misclassification, particulary frequent in asthma diagnoses.

These limitations aside, the case-control approach used in this study included in the model the time of occurrence of the events (admissions), even for the series of controls, and needed neither adjustments for complex cyclical trends, nor the introduction of smoothing parameters. Moreover, the PPR models allowed us to analyze the effects of the pollutants many days before the admission, without the problem of collinearity.

In conclusion, the results of this study are comparable with those of other studies published in the literature. In particular, the estimated effect of each pollutant in the two weeks before the admission was comparable with the effects shown in other Italian studies (28) and the risks estimated were comparable with those reported in the literature (24) (Figure 3).



* Results of our study.

Figure 3. Percentage of increased risk (and relative 95% confidence interval) of hospital admission for respiratory causes per 10 $\mu g/m^3$: comparison with other cities (24).

The approach used in this study appears to be useful for estimating the risk of respiratory illness admission associated with air pollution exposure and the results confirm that the effect occurs in the days before admission, underlining the need for further research in this area.

References

1. Her Majesty's Public Health Service. Mortality and morbidity during the London fog of December 1952. Report No. 95 on Public Health and Medical Subjects.
2. Martin AE, Bradley WH. Mortality, fog and atmospheric pollution: an investigation during the winter of 1958-1959. Mon Bull Minist Health Public Health Lab Serv 1960; 19: 56-72.
3. Martin AE. Mortality and morbidity statistics and air pollution. Proc R Soc Med 1964; 57 (Suppl): 969-975.
4. Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay M, Ferris BG Jr, Speizer FE. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993; 329: 1753-1759.
5. Pope CA 3rd, Thun MJ, Namboodiri M, Dockery DW, Evans JS, Speizer FE, Heath CW Jr. Particulate air pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 1995; 151: 669-674.
6. Barbone F, Bovenzi M, Cavallieri F, Stanta G. Air pollution and lung cancer in Trieste, Italy. Am J Epidemiol 1995; 141: 1161-1169.
7. Biggeri A, Barbone F, Lagazio C, Bovenzi M, Stanta G. Air pollution and lung cancer in Trieste, Italy: spatial analysis of risk as a function of distance from sources. Environ Health Perspect 1996; 104: 750-754.
8. Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 2002; 287: 1132-1141.
9. Bates DV, Sizto R. Air pollution and hospital admissions in Southern Ontario: the acid summer haze effect. Environ Res 1987; 43: 317-331.
10. Thurston GD, Ito K, Kinney PL, Lippmann M. A multi-year study of air pollution and respiratory hospital admissions in three New York State metropolitan areas: results for 1988 and 1989 summers. J Expo Anal Environ Epidemiol 1992; 2: 429-450.
11. Burnett RT, Dales RE, Raizanne ME, Krewski D, Summers PW, Roberts GR, Raad-Young M, Dann T, Brook J. Effects of low ambient level of ozone and sulfates on the frequency of respiratory admission to Ontario hospitals. Environ Res 1994; 65: 172-194.
12. Schwartz J. Air pollution and hospital admissions for the elderly in Birmingham, Alabama. Am J Epidemiol 1994; 139: 589-598.
13. Schwartz J. Air pollution and hospital admissions for respiratory disease. Epidemiology 1996; 7: 20-28.
14. Schwela D. Air pollution and health in urban areas. Rev Environ Health 2000;15:13-42.
15. Sunyer J, Sáez M, Murillo C, Castellsague J, Martinez F, Antó JM. Air pollution and emergency room admissions for chronic obstructive pulmonary disease: a 5-year study. Am J Epidemiol 1993; 137: 701-705.
16. Katsouyanni K, Zmirou D, Spix C. Short-term effects of urban air pollution on health: a European approach using epidemiologic time series data. The APHEA protocol. J Epidemiol Community Health 1996; 50 (Suppl. 1): S12-S18.
17. Katsouyanni K, Zmirou D, Spix C, Sunyer J, Schouten JP, Pönkä A, Anderson HR, Le Moullec Y, Wojtyniak B, Vigotti MA, et al. Short-term effects of urban air pollution on health: a European approach using epidemiologic time series data. The APHEA project: background, objectives, design. Eur Respir J 1995; 8: 1030-1038.
18. Spix C, Anderson HR, Schwartz J, Vigotti MA, LeTertre A, Vonk JM, Touloumi G, Balducci F, Piekarski T, Bacharova L, Tobias A, Pönkä A, Katsouyanni K. Short-term effects of air pollution on hospital admissions of respiratory diseases in Europe: a quantitative summary of APHEA study results. Air Pollution and Health: a European Approach. Arch Environ Health 1998; 54: 54-64.
19. Atkinson RW, Anderson HR, Sunyer J, Ayres J, Baccini M, Vonk JM, Boumghar A, Forastiere F, Forsberg B, Touloumi G, Schwartz J, Katsouyanni K. Acute effects of particulate air pollution on respiratory admissions: results from APHEA 2 project. Air Pollution and Health: a European Approach. Am J Respir Crit Care Med 2001; 164: 1860-1866.
20. Sunyer J, Atkinson R, Ballester F, Le Tertre A, Ayres JG, Forastiere F, Forsberg B, Vonk JM, Bisanti L, Anderson RH, Schwartz J, Katsouyanni K; APHEA 2 study. Respiratory effects of sulphur dioxide: a hierarchical multicity analysis in the APHEA 2 study. Occup Environ Med 2003; 60: 2e.
21. Fusco D, Forastiere F, Michelozzi P, Spadea T, Ostro B, Arcà M, Perucci CA. Air pollution and hospital admissions for respiratory conditions in Rome, Italy. Eur Respir J 2001; 17: 1143-1150.
22. Oftedal B, Nafstad P, Magnus P, Bjørkly S, Skrondal A. Traffic related air pollution and acute hospital admission for respiratory diseases in Drammen, Norway 1995-2000. Eur J Epidemiol 2003; 18: 671-675.
23. Buckeridge DL, Glazier R, Harvey BJ, Escobar M, Amrhein C, Frank J. Effect of motor vehicle emissions on respiratory health in an urban area. Environ Health Perspect 2002; 110: 293-300.
24. Biggeri A, Bellini P, Terracini B; Italian MISA Group. Meta-analysis of the Italian studies on short-term effects of air pollution. Epidemiol Prev 2001; 25 (Suppl. 2): 1-71.
25. Ciccone G, Faggiano F, Falasca P. SO2 air pollution and hospital admissions in Ravenna: a case-control study. Epidemiol Prev 1995; 19: 99-104. [Published in Italian].

26. Panella M, Palin L, Binotti M, Tommasini V, Di Stanislao F. The role of atmospheric nitrogen dioxide in the risk of hospital admission in the patient with respiratory pathology. Epidemiology. 2000 May; 11 (3): 367-8.
27. Migliaretti G, Cadum E, Migliore E, Cavallo F. Traffic air pollution and hospital admission for asthma: a case-control approach in a Turin (Italy) population. Int Arch Occup Environ Health 2005; 78: 164-169.
28. Biggeri A, Bellini P, Terracini B. Meta-analysis of the Italian studies on short-term effects of air pollution--MISA 1996-2002. Epidemiol Prev 2004; 28 (4-5 Suppl): 4-100. [Published in Italian].
29. World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Ninth Revision. Geneva; World Health Organization 1977.
30. Friedman JH, Stuetzle W. Projection pursuit regression. Journal of the American Statistical Association 1981; 76: 817-823.
31. World Health Organization. Acute effects on health of smog episodes. WHO Regional Publications, European Series (N.43). Geneva: WHO Regional Office for Europe, 1992.
32. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2005.
33. Costa G, Cardano M, Demaria M. Torino: Storie di salute in una grande città. City of Turin, Socio-economic Observatory, Statistical Unit, 1998.