Investigating Regional Differences in Short-Term Differences of Air Pollution on
Daily Mortality in the APHEA Project: A Sensitivity Analysis for Controlling
Long-Term Trends and Seasonality

Evangelia Samoli,1 Joel Schwartz,2 Bogdan Wojtyniak,3 Giota Touloumi,1 Claudia Spix,4 Frank Balducci,5 Sylvia
Medina,6 Giuseppe Rossi,7 Jordi Sunyer,8 Ljuba Bacharova,9 Hugh Ross Anderson10 and Klea Katsouyanni1

1Department of Hygiene and Epidemiology, National School of Public Health, Athens, Greece; 2Environmental Epidemiology
Program, Harvard School of Public Health, Boston, Massachusetts, USA; 3National Institute of Hygiene, Warsaw, Poland;
4GSF—Forschungszentrum für Umwelt und Gesundheit, Institute für Epidemiologie Postfach, Neuherberg, Germany; 5Faculté de
Medecine, Université de Grenoble, Department of Public Health, Grenoble, France; 6Observatoire Regional de la Santé, Paris, France;
7Institute of Clinical Physiology, National Research Council, Pisa, Italy; 8Institut Municipal D’Investigacion Medica, Barcelona, Spain;
9National Centre for Health Promotion, Bratislava, Slovakia; 10Department of Public Health Sciences, St. George’s Hospital Medical
School, London, United Kingdom

Short-term effects of air pollution on daily mortality in eight western and five central-eastern
European countries have been reported previously, as part of the APHEA project. One intriguing
finding was that the effects were lower in central-eastern European cities. The analysis used sinu-
soidal terms for seasonal control and polynomial terms for meteorologic variables, but this is a
more rigid approach than the currently accepted method, which uses generalized additive models
(GAM). We therefore reanalyzed the original data to examine the sensitivity of the results to the
statistical model. The data were identical to those used in the earlier analyses. The outcome was
the daily total number of deaths, and the pollutants analyzed were black smoke (BS) and sulfur
dioxide (SO2). The analyses were restricted to days with pollutant concentration < 200 µg/m³ and
< 150 µg/m³, respectively. We used Poisson regression in a GAM model, and combined individual
city regression coefficients using fixed and random-effect models. An increase in BS by 50 µg/m³
was associated with a 2.2% and 3.1% increase in mortality when analysis was restricted to days
< 200 µg/m³ and < 150 µg/m³, respectively. The corresponding figures were 5.0% and 5.6% for a
similar increase in SO2. These estimates are larger than the ones published previously: by 69% for
BS and 55% for SO2. The increase occurred only in central-eastern European cities. The ratio of
western to central-eastern cities for estimates was reduced to 1.3 for BS (previously 4.8) and 2.6
for SO2 (previously 4.4). We conclude that part of the heterogeneity in the estimates of air pollution
effects between western and central-eastern cities reported in previous publications was
caused by the statistical approach used and the inclusion of days with pollutant levels above 150
µg/m³. However, these results must be investigated further. Key words: air pollution, black
smoke, generalized additive models, mortality, Poisson regression, sensitivity analysis, sulfur diox-

In the last decade an extensive body of epide-
nemiologic literature, initially mainly from
North America, has reported associations
between routinely occurring air pollution
conzentations and daily health outcomes
(1–3). Short-term effects of air pollution on
daily deaths have been investigated in a large
European multicenter study, the Air Pollu-
ion and Health: A European Approach
(APHEA) project, which included data from 15
cities, including 5 in central-eastern
Europe (4 Polish and 1 Slovakian). The
central-eastern European cities contributed
exposure data on sulfur dioxide and black
smoke levels only. In the original APHEA
project, data from each individual city were
analyzed using a standardized protocol (4)
based on sinusoidal terms for seasonal con-
trol and polynomial terms for meteorologic
variables. One intriguing finding was that
the effects were lower (although still statisti-
cally significant in most instances) in cen-
tral-eastern European cities (5,6) than in
western European cities. In the published
findings of the study, the authors postulated
that “the model for seasonal control may fit
the data less well in central-eastern cities
because of a higher and more variable rate of
respiratory illness” (5). In addition, the cen-
tral-eastern European cities had higher con-
centrations of air pollution than the Wester-
n cities.

Early epidemiologic studies used simple
techniques to control for season and
weather, such as indicator variables for sea-
sons and hot days and linear terms for
weather factors (7,8). Sinusoidal terms for
seasonal control were then introduced to
provide a better fit (9,10). More recent studies
used generalized additive models (GAM),
which allow nonparametric smooth func-
tions to control for season and weather
(11,12), and this method is rapidly becom-
ing standard practice, combined with sensi-
itivity analyses (13–15).

The GAM approach requires specialized
software that was not in widespread use at
the time the APHEA project started (4,12).

And because studies in North America had
reported little sensitivity of the air pollution
results to the method of seasonal control
(16,17), the APHEA group decided to use
sinusoidal terms for seasonal control and
polynomial terms for weather (4). However,
because of the heterogeneity of the findings
of the APHEA project described above, the
APHEA group has now undertaken a sensi-
tivity analysis of the results using GAM, to
test the adequacy of seasonal control and to
provide a basis for comparison with the
results of the new APHEA 2 project that is
now in progress. APHEA 2 will use GAM
and will involve more than 30 European
cities, including about 10 from central-east-
ern Europe.

The APHEA studies, like earlier studies,
found in most instances a nonlinear relation
between air pollution and daily deaths. The
dose-response curve was roughly logarith-
omic, with lower slopes at higher levels.
Fitting a linear model across a range of pol-
lution where the relationship is nonlinear
may also account for some of the differences
between cities in eastern and western
Europe. To test this, we have also examined
the sensitivity to that change. To facilitate
the combining of slopes and provide a slope
that was meaningful in the range of exposure
where standards are being considered,
APHEA fit linear pollution terms for days
with concentrations below 200 µg/m³.
APHEA 2 is planning to restrict analysis to
days below 150 µg/m³ for SO2 and particu-
late matter (PM). This will mostly affect
the central-eastern European cities, which have
higher pollution levels.

Address correspondence to K. Katsouyanni,
Department of Hygiene and Epidemiology,
University of Athens, Medical School, 75 Mikras
Asias Str., Athens 115 27, Greece. Telephone:
301-7719725. Fax: 301-7704225. E-mail:
kkatsoy@cc.uoa.gr

The APHEA project was supported by the
European Commission, DG XII, Environment
1991-94 Programme (contract number EV5V
CT92-0202). Received 8 February 2000; accepted
14 November 2000.
Methods

The data used in this analysis are from seven western European cities—Athens, Barcelona, Cologne, London, Lyon, Milan, and Paris—and five central-eastern European cities—Bratislava, Cracow, Lodz, Poznan, and Wroclaw. The pollutants studied were sulfur dioxide (SO₂; 24-hr, provided by all cities) and black smoke (BS; provided by four western and four central-eastern cities). Gravimetric PM data were provided by Lyon and Paris (PM₁₃), Cologne (PM₂.5), Barcelona, Bratislava, and Milan (total suspended particulates). The populations in these cities range from 400,000 to >7 million. The daily number of deaths from all causes, excluding deaths from external causes, was the health outcome. The data covered at least 5 consecutive years for each city within the years 1980–1992. Details about the data have been published elsewhere (5,6,18). The analyses were restricted to days when the levels were <200 µg/m³ or 150 µg/m³ for SO₂ and BS, respectively, because in these lower ranges roughly 40% associations with the logarithm of the expected mortality are observed.

Restricting analyses to days with <200 µg/m³ meant that <5% of the available days were excluded, except for Cracow, for which 5.9% of the days were excluded when SO₂ was analyzed and 15.4% of the days when BS was analyzed. Restricting analyses to days with levels <150 µg/m³ resulted in exclusion of 5.6, 6.8, and 14.9% of days for Lodz, Poznan, and Cracow, respectively, for the SO₂ analyses and of 9.3, 10.2, 8.5, and 24.6% of days for Athens, Lodz, Wroclaw, and Cracow, respectively, for the BS analyses; the percentage of days excluded in all other cities was <5% for both pollutants.

We investigated pollution–mortality associations using Poisson regression in a GAM (19,20). This model allowed us to include nonparametric smooth functions to model the potential nonlinear dependence of daily admissions on weather and season. It assumes that

\[ \log(E(Y)) = \beta_0 + S_1(X_1) + \ldots + S_p(X_p), \]  

where \( Y \) is the daily count of admissions, \( E(Y) \) is the expected value of that count, the \( X_i \) are the covariates, and the \( S_i \) are the smooth functions. We chose loess (21), a moving regression smoother. This is a generalization of a weighted moving average, and it estimates a smooth function by fitting a weighted regression within a moving window (or fraction of the data) centered about each value of the predictor variable. The weights are close to one for the central third of the window and decline to zero rapidly outside that range. Outside of the window, the weights are all zero. The covariates we controlled for included temperature, relative humidity, day of the week, epidemic periods, and holidays.

Choice of smoothing window. The critical choice in a nonparametric smoother is the size of the smoothing window. Here we distinguished between weather variables, which we believe are causally connected to deaths, and seasonal control. For temperature and relative humidity we chose the span that minimizes Akaike’s Information Criterion (AIC) (22). AIC trades off the improvement in fit from using smaller windows against the increase in degrees of freedom used. It is roughly equivalent to minimizing the deviance on a validation data set. Hence, underfitting is penalized (because the deviance in the validation set is higher) and overfitting is also penalized (because if we fit noise in the observed data, that pattern will not be present in the validation data set).

The choice of window size for time is a different question. Day of study is not thought to be a causal variable. Rather, we know that our regression excludes many risk factors for mortality, such as smoking and diet. These are not confounders if they are not correlated with air pollution. It may reasonably be assumed that their daily or short-term variation is not correlated with air pollution levels. However, if there are long-term trends or seasonal patterns in these omitted factors, then they may be correlated with air pollution, because it, too, varies seasonally and has time trends. In this sense, time is used as a proxy for any outcome predictors not included in the model, which vary over time as described. Hence we remove long-term trends and seasonal patterns from the data, with a smooth function of time, to guard against this confounding by omitted variables.

Our goal is not to remove all pattern from the fluctuations in daily deaths, but rather to remove all fluctuations that are seasonal or longer. AIC is thus not the appropriate criterion (12). Rather, a window between 80 and 200 was decided a priori. Smooth plots with windows of 80–200 days fit the basic seasonal patterns. They allowed the winter peak in mortality to vary from season to season in location and height and were short enough to include double peaks in mortality in winters with two serious respiratory epidemics. Hence they appear basically adequate to the task. The use of windows of less than 80 days tended to fit much shorter duration patterns in the data, such as fluctuations of 1–2 weeks, which could be caused by air pollution. Within that range, we chose the span for each city that minimized the autocorrelation of the residuals. We used this approach because each death is an independent event, and autocorrelation in residuals indicates that there are omitted time-dependent covariates whose variation may confound air pollution. If the autocorrelation is removed, remaining variation in omitted covariates has no systematic temporal pattern, so confounding is less likely. In contrast, overfiltering can produce high frequency “ringing” in the data that induces autocorrelation (23). “Ringing” refers to the tendency of high-pass filters to produce high-frequency distortion. The tendency to induce negative autocorrelation by excessive seasonal control has been noted by Diggle (24). This can distort the association...
between air pollution and deaths. Thus, minimizing the autocorrelation within a class of models that remove seasonal patterns is a reasonable objective. In practice we first minimized the autocorrelation of the seasonal model and then minimized the AIC for the weather terms, holding the seasonal model fixed. We then reexamined the seasonal span and minimized autocorrelation.

We controlled for day-of-the-week effects, holidays, and epidemics using dummy variables. We used robust regression (25) to reduce the effect of any extreme observations on the regression results.

Choice of lag. Air pollution and weather may have an immediate effect on daily deaths but may also produce delayed effects. The APHEA I protocol called for assessment of the association with weather and pollutant terms on the same day as the death and on the days immediately prior (4). The lag of each variable that best fit the data was chosen in that study. To maintain comparability, we used the same lags. We compare results using the same single-day lag of each air pollutant, chosen in the APHEA study. Most published studies that examined the question have found that the average of several days’ pollution correlate better with mortality than a single day’s exposure. To account for this while maintaining comparability across cities, we decided to also analyze the average of lags 0 and 1 for all cities using the new methodology. As in APHEA I, single-pollutant models were fitted because the correlations between SO$_2$ and the particle measures were too high to allow stable estimates in two-pollutant models. The analysis was done using S-plus software (MathSoft Inc., Cambridge, MA, USA).

Combining results over cities. Once models were fitted in each location, we summarized the results over all locations using inverse variance weighting. For the fixed-effects meta-analyses, the estimated overall effect was a weighted average with weights taken to be the inverse of the square of the standard errors of the pollution coefficient. We computed separate summaries for eastern Europe, western Europe, and for all of the cities. We examined heterogeneity by computing chi-square statistics (26). When there was significant heterogeneity, we also computed pooled coefficients using random-effects models. These estimated the overall effect as a weighted average, with weights equal to the inverse of the sum of the square of the standard error plus a random variance component. The random variance component was estimated using the method of moments (26).

Results

The new more flexible model produced considerable changes in the estimated seasonal and weather effects, particularly for the central-eastern cities. Figure 1 shows the estimated seasonal pattern in Lodz using the parametric model and the nonparametric smooth. The parametric model has the same difference between summer and winter in each year, has a shoulder in the fall of each year, and has a double peak of mortality in each year. The nature of the trigonometric functions forces the double peak to occur either in each year or not at all. The nonparametric model allows the winter-to-summer difference to change from year to year, which it clearly did in this case. It also shows a double peak of wintertime mortality only in some years.

Figure 2 shows the estimated individual city and pooled relative risks of mortality associated with an increase of 50 µg/m$^3$ in sulfur dioxide concentration, restricted to days < 200 µg/m$^3$, using the old and new methodology for seasonal control. It can be seen that most estimated relative risks for individual cities have increased. The pooled estimated increase in daily mortality is now larger by 55% over all cities and has changed proportionally more in the central-eastern European cities, where, however, its magnitude still remains about half what is seen in the west (Table 1). The ratio of the pooled estimated increase in daily mortality of western to central-eastern European cities was 4.4 with the old methodology and is 2.6 for the best 1-day lag and 2.4 for the average of lags 0 and 1. There is still statistically significant heterogeneity between the estimates of the individual cities, and there remains heterogeneity in the western cities, mainly due to the higher effect estimates in Lyons, Athens, and Barcelona. Heterogeneity was also introduced in the data for central-eastern European cities because of the relatively low effect estimates in Bratislava.

Figure 3 shows the estimated individual city and pooled relative risks of mortality associated with an increase of 50 µg/m$^3$ in SO$_2$ concentration, restricted to days < 200 µg/m$^3$, using the old and the new methodology. The individual-city relative risks for western cities are lower with GAM for three cities and higher for the fourth. In central-eastern cities there is a substantial increase in the relative risks in all four cities, although in Poznan the effect is still not statistically significant. The pooled estimates are slightly higher in the western cities but have substantially increased for the central-eastern, although...
they remain lower than in the west (Table 2). The ratio of western to central-eastern cities increases in daily mortality associated with 50 µg/m³ change in BS levels was 4.8 with the old methodology and 1.3 for the same-day lag and 1.8 for the models using average 0 and 1 lags.

Table 3 shows the estimated effects for 

SO₂ and BS when the analysis is restricted to days with pollution concentrations < 150 µg/m³. In the western cities, which had few days > 150 µg/m³, the effect estimates are little changed. In the central-eastern cities, which had more days with concentrations between 150 and 200 µg/m³, the effect estimates increase for both pollutants. For BS there is no longer any difference between the estimates in the east and the west. For SO₂, the percent increase in effect size was larger in the central-eastern cities than in the western cities after the restriction, but there remained a difference in the overall effect.

**Discussion**

We have presented a more sophisticated analytic method for epidemiologic time-series studies applied to data previously analyzed with a more rigid approach. The GAM method allows more flexibility in the control of confounders, either identified (like temperature) or unidentified. Specifically, it allows better control of time trends and seasonality, which refer to patterns covering longer time periods (12).

The GAMs applied in the sensitivity analysis presented here generally led to increases in the estimated pooled relative risks of total mortality associated with higher concentrations of sulfur dioxide and black smoke in the ambient air. The changes were smaller in the Western European cities. For a 50 µg/m³ increase in SO₂, the increase in mortality in western cities was 3.5% using sinusoidal terms for seasonality (old method) and 5.0% using a GAM. The corresponding figures for a similar change in BS levels were 2.9% and 3.2%, respectively, which we view as essentially identical. In central-eastern European cities the estimated change in daily mortality increased proportionally more. For a difference of 50 µg/m³ in SO₂ concentrations, the estimated increase in mortality was 0.8% using the old methodology and 1.9% using the new, and for BS it was 0.6% and became 1.9%. However, the estimates in the central-eastern European cities, although now closer to the ones estimated for the western European cities, remain lower by about 50% for both pollutants.

Restricting the analysis to days with concentrations < 150 µg/m³ further reduced the differences between the western European and central-eastern European cities. For BS there was practically no difference between the effect size estimates between the two regions. For SO₂, these factors slightly reduced the regional differences in the estimates, which remained lower in the central-eastern cities by about 50%. Again, this restriction had little effect in western Europe. SO₂ may represent different mixtures of air pollution in western and central-eastern cities, and this may explain the persistent difference. This confirms our hypothesis that the previously observed differences could be explained partly by poorer seasonal control and nonlinearities in the dose–response relationship at higher concentrations.

The original APHEA paper (5) also reported an association between gravimetrically measured airborne particles (PM and total suspended particles) and daily deaths. We have not emphasized that result because the diverse ways in which particles were measured in the cities make comparisons difficult. However, it is worth noting that for PM₁₀ concentrations, the effect size estimates increased, and the estimated increase for 50 µg/m³ of PM₁₀ became 3.3% (95% confidence interval, 2.6–4.1) using the GAM model. This is similar to the results reported in North America (2).

We conclude that part of the heterogeneity in the air pollution estimates between central-eastern and western European cities reported in previous publications (5) was caused by inadequate control of seasonality by the sinusoidal terms and inclusion of concentrations where the dose–response relationship became nonlinear. However, heterogeneity remains, and in the context of the present study the limited number of cities does not allow more insight beyond previous results (5). This heterogeneity will be investigated as part of the current APHEA 2 project. Furthermore, the heterogeneity that has been observed across all of Europe remains statistically significant, as well as within western European cities. The lack of heterogeneity for BS estimates in the central-eastern European cities may well be explained by the fact that all are Polish cities, which probably share common characteristics, whereas the western European cities belong to four different countries. In the APHEA 2 project more than 30 cities will be
analyzed, more than 10 of which belong to 6 central-eastern European countries. This will provide a better opportunity to investigate the influence of different seasonality patterns and other effect modifiers.

**REFERENCES AND NOTES**