pattern between particulate matter with aerodynamic diameters < 10 μm (PM_{10}) and cause-specific mortality in a multicity hierarchical model. Only one study (Schwartz 2000c) has looked at mortality displacement with regard to cause-specific deaths.

Some studies found an association between airborne particles and mortality by age groups (Bremner et al. 1999; Goldberg et al. 2000; Ostro et al. 1999a; Ponka et al. 1998; Verhoeff et al. 1996), but again, none analyzed the distribution of the effect over time since exposure and whether the mortality displacement is short term.

In our previous study of the APHEA-2 (Air Pollution and Health: A European Approach) project (Zanobetti et al. 2002), we examined mortality displacement between air pollution and all daily deaths. The goal of that analysis was to estimate the dependence of daily deaths on day t on PM_{10} on that day, and up to the previous 40 days using an unconstrained distributed lag model. As a sensitivity analysis, in that study we used a fourth-degree polynomial and a cubic distributed lag model, and all the models showed similar results.

In the present study, within the APHEA-2 project, we examined mortality displacement by age, and we looked at the differences between cardiovascular and respiratory deaths compared with all causes of deaths. Following the same approach as in our previous analysis, we applied a fourth-degree polynomial and unconstrained distributed lag models in a multicity hierarchical framework.

Data and Methods

Health data. The APHEA-2 project was a multicenter study of 30 cities across Europe and adjacent countries (e.g., Istanbul, Turkey, and Tel Aviv, Israel). The goal of the APHEA-2 project was to examine the effects of air pollution on daily mortality. Details on APHEA-2 data have been reported previously (Katsouyanni et al. 2001).

To maximize the power of the study, we chose the largest cities in the project, subject to the constraint that only one city be chosen in each country. The 10 cities selected were Athens, Greece; Budapest, Hungary; Lodz, Poland; London, United Kingdom; Madrid, Spain; Paris, France; Prague, Czech Republic; Rome, Italy; Stockholm, Sweden; and Tel Aviv. Together, they comprise a population of about 28 million people, which is two-thirds of the population in the full study, and represents Northern Europe, Central Europe, and the Mediterranean region.

The data that we analyzed in this study included daily counts of all-cause mortality, excluding deaths from external causes [*International Classification of Disease*, *9th Revision* (*ICD-9*) code > 800), daily counts of deaths from cardiovascular disease (CVD) (*ICD-9* code 390–429), and daily counts of respiratory mortality (*ICD-9* code 460–519). In all cities, the mortality counts were also divided in the following age groups: 15–64, 65–74, and ≥ 75 years.

The years of study were 1990 through 1997; however, mortality data in most cities were available only through 1995 or 1996.

Air pollution data were available as daily averages of the monitoring stations in each city. Particulate matter was measured as PM_{10} in four cities, as PM_{13} in Paris and PM_{15} in Rome. The Paris data were assumed to be equivalent to PM_{10} in this study. In Rome, Athens, Lodz, and Budapest, the data were converted from other measurements into PM_{10} using city-specific conversion factors. Further details have been published previously (Katsouyanni et al. 2001; Zanobetti et al. 2002).

We first analyzed respiratory and heart mortality; then, we examined deaths from all causes and stratified respiratory and cardiovascular deaths by age groups.

Methods. In the first step of the analysis, generalized additive regression models (Hastie and Tibshirani 1990) were fit in each of the 10 cities, and the models were built following the APHEA-2 methodology. Details on the models for all-cause mortality have been published (Katsouyanni et al. 2001).

Separate models were built in each city because of the substantial variability in seasonal patterns and weather between, for example, Stockholm and Tel Aviv.

In the generalized additive model, the outcome is assumed to depend on a sum of nonparametric smooth functions for each variable. This allows us to better model the nonlinear dependence of daily deaths on weather and season. The model is of the form

$$\log[E(Y_t)] = \alpha_0 + S_1(X_1) + S_i(X_i) + \ldots + S_p(X_p),$$

where $E(Y_t)$ is the expected value of the daily count of deaths Y_t, and S_i is the smooth function of the covariate X_i. The covariates we examined were weather, influenza epidemics, holidays, and day of the week. The locally weighted running-line smoother (Cleveland and Devlin 1988) was chosen to estimate the smooth function.

All models controlled for temperature and relative humidity on the same day using nonparametric smooth functions. In addition, we examined nonparametric functions and linear weather variables on the previous day or up to 3 previous days or the average of a few days. We chose the variables that minimized Akaike's Information Criterion (Akaike 1973) for the model. This approach has been used and discussed previously (Katsouyanni et al. 2001; Schwartz 1999; Schwartz et al. 1996).

To remove seasonal and long-term fluctuations, we used a smooth function of time. Seasonal patterns are controlled because there are unmeasured predictors of death, such as diet, which have long-term trends over time and vary seasonally. Because air pollution also shows seasonal variations, and often long-term trends, this creates a potential for confounding. Day-to-day fluctuations in diet are unlikely to be correlated with air pollution.

To model seasonality, we chose the smoothing parameter that minimized the sum of the autocorrelation of the residuals while removing seasonal patterns.

In some models, it was necessary to introduce autoregressive terms to eliminate the remaining serial correlation from the residuals (Brumback et al. 2000). This approach has been used in a number of recent studies (Rossi et al. 1999; Schwartz 1999, 2000a).

Recently several authors have pointed out that there are problems with the implementation of generalized additive models in current statistical packages (Dominici et al. 2002; Ramsay et al. 2003). One is the lax default convergence criterion in generalized additive models in S-PLUS software (MathSoft, Inc., Seattle, WA). A more serious problem is that this program does not properly estimate standard errors. To deal with the underestimation of the parameters' variances, we calculated standard errors for the estimated regression coefficients using the nonparametric bootstrap (Efron and Tibshirani 1993). In particular, we generated 250 bootstrap samples each consisting of n values drawn with replacement from the observed data, where n was the number of observations in each city. The model of interest was run for each generated data set, and the sample standard deviation of the resulting regression coefficients serves as the bootstrap standard error.

We examined the dependence of daily deaths on PM_{10} of that day and of preceding days, up to the previous 40 days, using first an unconstrained distributed lag model and then a fourth-degree polynomial distributed lag model.

If the pollution-related deaths are only being advanced by a few days to a few weeks, we would see this "harvesting" effect expressed by negative associations between air pollution and deaths several days to weeks afterward. The effect of air pollution, net of any such short-term rebound up to 40 days, is the sum of the effect estimates for all 41 days.

For Poisson regression, the unconstrained distributed lag model can be written as

$$\mathrm{Log}\left[E(Y_t)\right] = \alpha + \mathrm{covariates} + \beta_0 Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} \quad [1]$$

The unconstrained distributed lag model is too noisy to provide any information about

the shape of the effect versus lag, but it does give an unbiased estimate of the overall effect, computed as the sum of the β_j. To be able to have an idea of the shape of the curve, it is preferable to constrain the coefficients to vary smoothly with lag number (Almon 1965). A polynomial distributed lag constrains the β_j to follow a polynomial pattern in the lag number; that is,

$$\beta_j = \sum_{k=0}^{d} \eta_k j^k, \quad \text{for } j = 0 \ldots q, \qquad [2]$$

where j is the number of lags and d is the degree of the polynomial.

Too much constraint risks bias, producing a distorted shape, whereas too little constraint produces estimates that are too noisy to be informative.

To fit a fourth-degree polynomial distributed lag model with 40 days of delay, in equation 2 we would have $d = 4$ and $j = 0, \ldots, 40$.

In the first stage of the analysis, a separate distributed lag model was fit for each of the 10 cities. The coefficients $\hat{\beta}_j$ by lag for the distributed lag model were obtained from equation 2, and because we used the bootstrap method, we computed the sum of the effect estimates for all 41 days in each city. In the second stage, we combined the city-specific results using inverse variance weighted averages.

In our primary analysis we used an unconstrained distributed lag, because it makes the fewest assumptions. We used the polynomial distributed lag model as a sensitivity analysis because this gives more stable results and also allows us to examine the pattern of the mortality response over time.

To verify whether the longer lags really contribute to the overall pollution effect, we also analyzed the unconstrained distributed lag model with 20 and 30 days of delay.

Results

Table 1 shows, for the 10 cities, the mean and standard deviation of the daily number of total, cardiovascular, and respiratory deaths, in total and for the three age groups. It also shows the mean and standard deviation of the environmental variables. The mean of daily deaths for 15- to 64-year-olds is generally very low when subdivided by cause. For respiratory mortality, the second age group is also small; therefore, we decided that the cause-specific analysis by age could be done for the two older groups for cardiovascular mortality, and in the oldest age group for respiratory mortality.

Table 2 shows the percentage increase for the combined effect estimates across all of the 10 cities of PM_{10} [per 10 μg/m^3 and its 95% confidence interval (95% CI)] for the fourth-degree polynomial (sum of the β_j per 10 μg/m^3), and the unconstrained distributed lag models when stratified by cause of death.

Figures 1 and 2 show plots of the residuals from the unconstrained distributed lag model for Athens and London, respectively, with no residual seasonality in the models.

We found a 4.2% increase (95% CI, 1.08–7.42) in respiratory deaths for a 10 μg/m^3 increase in PM_{10} concentration using the unconstrained distributed lag model. Similar results, but with tighter confidence intervals, were found with the polynomial distributed lag models. In contrast, the mean of PM_{10} on the same and previous day was associated with only

Table 1. Study period, population, and mean ± SD of the number of daily deaths and the environmental variables in the 10 cities.

City/age group (years)	Years of study	Mortality			PM_{10} (mg/m^3)	Temperature	Humidity
		Total	CVD	Respiratory			
Athens	1992–1996	72.9 ± 13.2	35.9 ± 8.7	4.6 ± 2.4	42.7 ± 12.9	17.8 ± 7.4	61.7 ± 13.6
15–64		8.3 ± 3	2.9 ± 1.7	0.4 ± 0.6			
65–74		21.3 ± 5.2	9.3 ± 3.4	1.1 ± 1.1			
≥ 75		42.4 ± 9.6	23.7 ± 6.7	3.2 ± 1.9			
Budapest	1992–1995	80 ± 11.6	40.1 ± 8	3.1 ± 2	41 ± 9.1	12.8 ± 8.8	70.1 ± 12.6
15–64		22.1 ± 4.7	6.8 ± 2.7	0.6 ± 0.8			
65–74		19.5 ± 4.6	9.5 ± 3.2	0.7 ± 0.8			
≥ 75		37.6 ± 7.9	23.7 ± 6	1.7 ± 1.4			
Lodz	1990–1996	29.5 ± 6.3	16.8 ± 4.6	1.2 ± 1.1	53.5 ± 15.5	8.4 ± 8.4	79 ± 12.4
15–64		8.9 ± 3.2	3.7 ± 2	0.3 ± 0.5			
65–74		8.1 ± 3	4.5 ± 2.2	0.3 ± 0.6			
≥ 75		12.2 ± 3.9	8.7 ± 3.2	0.5 ± 0.8			
London	1992–1996	168.5 ± 25.2	71.4 ± 12.9	29 ± 11	28.8 ± 13.7	11.8 ± 5.4	69.3 ± 11.3
15–64		28.8 ± 5.7	9.5 ± 3.2	2.4 ± 1.7			
65–74		37 ± 7.4	15.7 ± 4.5	4.6 ± 2.6			
≥ 75		101.7 ± 18.5	46.1 ± 9.6	21.8 ± 8.9			
Madrid	1992–1995	60.8 ± 11.1	21.7 ± 6.1	6.1 ± 3.1	37.8 ± 17.7	14.5 ± 7.4	61.8 ± 16.7
15–64		13.7 ± 3.9	3.1 ± 1.8	0.7 ± 0.9			
65–74		13 ± 3.8	4 ± 2.1	1.1 ± 1.1			
≥75		33.4 ± 7.7	14.6 ± 4.6	4.2 ± 2.5			
Paris	1992–1996	123.3 ± 15.7	37.9 ± 7.5	9 ± 3.8	22.5 ± 11.5	12.1 ± 6.5	75.6 ± 12.5
15–64		30.7 ± 5.7	4.1 ± 2	0.9 ± 0.9			
65–74		20.8 ± 5	5.4 ± 2.5	1.2 ± 1.1			
≥ 75		69.9 ± 11.8	28.3 ± 6.4	6.9 ± 3.2			
Prague	1992–1995	38.2 ± 7.2	22.2 ± 5.2	1.4 ± 1.2	76.2 ± 45.7	11 ± 8	69.4 ± 14.1
15–64		7.5 ± 3	2.8 ± 1.7	0.3 ± 0.5			
65–74		10.2 ± 3.2	5.3 ± 2.3	0.4 ± 0.6			
≥ 75		20.1 ± 5.1	14.1 ± 4.2	0.7 ± 0.9			
Rome	1992–1996	56.2 ± 10.4	23.2 ± 6.4	3.1 ± 2	58.7 ± 17.4	16.8 ± 6.7	61.6 ± 11.9
15–64		11.3 ± 3.6	2.7 ± 1.7	0.2 ± 0.5			
65–74		13.4 ± 4	4.5 ± 2.3	0.6 ± 0.8			
≥ 75		30.9 ± 7.5	16 ± 5.1	2.2 ± 1.7			
Stockholm	1994–1996	28.9 ± 6.1	14 ± 4.1	2.5 ± 1.7	15.5 ± 7.9	7.7 ± 8.1	71.4 ± 15.8
15–64		3.7 ± 2	1 ± 1.1	0.1 ± 0.4			
65–74		5.3 ± 2.4	2.2 ± 1.6	0.3 ± 0.6			
≥ 75		19.7 ± 5.1	10.7 ± 3.5	2 ± 1.5			
Tel Aviv	1993–1996	27.4 ± 6.3	11.5 ± 4.1	1.4 ± 1.4	50.3 ± 57.5	20.6 ± 5.4	65.6 ± 11
15–64		4 ± 2	1 ± 1.1	0.1 ± 0.4			
65–74		5.8 ± 2.5	2.3 ± 1.5	0.2 ± 0.5			
≥ 75		17.2 ± 4.9	8.2 ± 3.3	1 ± 1.2			