Europe[6,7] and the Harvard Six Cities project in the United States.[8] Major interest has been focused on potential factors that put people at increased risk of PM-related mortality and on the characteristics and sources of particles that affect their toxicity, as well as on the potential confounding by various other air pollutants. The impact of the above results has been important for the revisions of air-quality standards in the United States and Europe.[9,10]

The first multicenter studies were not based on extensive networks and were not able to explore confounding and effect modification satisfactorily. Also, in Europe the APHEA project included limited gravimetric measurement of particles. New multicenter studies include the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) study in the United States[11] and the APHEA2 study in Europe. These studies have tried to assess the consistency of the associations and to address questions of sensitive subpopulations, particle characteristics, and confounding.

We report here the results of the APHEA2 project on short-term effects of ambient particles on all-age and elderly all-cause mortality with emphasis on confounding factors and effect modifiers. In APHEA2 an extended database is used that includes more (29) cities and more extensive exposure data than the older APHEA project.[6] This database allows a more comprehensive and structured approach at the second stage of the analysis, in which we explore the role of effect modifiers in explaining heterogeneity.

Subjects and Methods

DATA

Location, Outcome, and Exposure

Table 1 shows the descriptive data from the 29 cities included in this analysis. The study period was longer than 5 years (1,826 days) for most cities. The total population in all cities was more than 43 million. We excluded 1 of the 30 cities originally providing data (Bucharest, Romania) because of insufficient days of PM concentration data (missing values 37%). The mean daily total number of deaths (excluding deaths from external causes, *International Classification of Diseases* code ≥800) ranged from 6 to 169 and that for the elderly (>65 years) ranged from 4 to 139. The median levels of black smoke (BS) and PM less than 10 μm in aerodynamic diameter (PM_{10}) concentrations (average of 2 consecutive days) ranged from 9 to 64 and from 14 to 166 $\mu g/m^3$, respectively. BS levels represent concentrations of black particles with an aerodynamic diameter <4.5 μm.[5] These measurements have a long history in Europe, and although standards for BS have been replaced recently by those for PM_{10},[12] the results are dis-

TABLE 1. Descriptive Data on the Study Period, Population, Exposure (PM_{10} and Black Smoke), Outcome (Daily Number of Deaths), and Levels of Other Pollutants (Exposure and Levels of Other Pollutants in $\mu g/m^3$)

Cities	Study Period	Population × 1000	No. of Deaths per Day		PM_{10} (24 Hour) Percentiles		Black Smoke (24 Hour) Percentiles		SO_2 (24 Hour) Percentiles		O_3 (1 Hour) Percentiles		NO_2 (24 Hour) Percentiles	
			Total	>65 Years	50	90	50	90	50	90	50	90	50	90
Athens	1/92–12/96	3,073	73	64	40*	59	64	122	46	89	83	135	74	114
Barcelona	1/91–12/96	1,644	40	32	60	95	39	64	12	14	71	112	69	97
Basel	1/90–12/95	360	9	8	28*	55			9	25	62	117	38	58
Bilbao	4/92–3/96	667	15	11			23	39	23	39			49	64
Birmingham	1/92–12/96	2,300	61	50	21	40	11	22	19	39	56	79	46	65
Budapest	1/92–12/95	1,931	80	57	40*	52			39	59	82	132	76	113
Cracow	1/90–12/96	746	18	13	54*	86	36	101	49	100			44	80
Dublin	1/90–12/96	482	13	10			10	26	21	34				
Erfurt	1/91–12/95	216	6		48	98			26	133	71	132	35	65
Geneva	1/90–12/95	317	6	4	33*	71			9	24	63	124	45	65
Helsinki	1/93–12/96	828	18	14	23*	49			6	16	57	83	33	50
Ljubljana	1/92–12/96	322	7	5			13	42	27	69	71	145	46	70
Lodz	1/90–12/96	828	30	20			30	77	19	56			39	59
London	1/92–12/96	6,905	169	139	25	46	11	22	22	36	43	71	61	86
Lyon	1/93–12/97	416	9	7	39	63			23	42	61	121	63	87
Madrid	1/92–12/95	3,012	61	46	33	59			26	75	52	98	70	107
Marseille	1/90–12/95	855	22	18			34	56	23	38			71	99
Milan	1/90–12/96	1,343	29	23	47*	88			20	82	38	119	94	141
Paris	1/91–12/96	6,700	124	91	22	46	21	45	15	34	38	84	53	74
Poznan	1/90–12/96	582	17	12			23	76	23	72			47	73
Prague	2/92–12/96	1,213	38	30	66	124			36	92	78	144	58	86
Rome	1/92–12/96	2,775	56	44	57*	81			11	23	41	93	88	115
Stockholm	1/90–12/96	1,126	30	25	14	27			4	10	63	91	26	37
Tel Aviv	1/91–12/96	1,141	27	22	43	75			19	39	36	56	70	156
Teplice	1/90–12/97	625	18	13	42	83			46	117	52	103	32	47
Torino	1/90–12/96	926	21	17	65*	129			23	71	88	159	76	118
Valencia	1/94–12/96	753	16	14			40	70	25	38	59	86	66	101
Wroclaw	1/90–12/96	643	15	10			33	97	21	62			27	39
Zurich	1/90–12/95	540	13	10	28*	54			10	28	62	123	40	59

PM_{10} = particulate matter less than 10 μm in aerodynamic diameter; SO_2 = sulfur dioxide; C_3 = ozone; NO_2 = nitrogen dioxide.
* PM_{10} values were estimated using a regression model relating collocated PM_{10} measurements to the black smoke or total suspended particles.

played here both for reasons of continuity and because there is evidence that BS exposure is more relevant to health effects than PM_{10}.[13,14] BS is a better marker of primary combustion products and small particles.[15] Because domestic or industrial burning of coal is minimal in most of the cities studied, BS is more specific for traffic-related particles than PM_{10} and provides a means of addressing the question of particle composition. Sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and ozone (O_3) concentrations are also shown. Measurements of air pollutants were provided by monitoring networks established in each town. European Union (EU) legislation regulates the methods of air pollutant measurements,[12] and recently most Central-Eastern European countries (which includes all cities in the former Communist countries and which are not EU members) have tried to comply with this regulation. Nevertheless, the recent EU daughter directive for PM_{10} measurement[12] (replacing the older one for BS) had not been applied during the time periods studied here and, as a result, there is variability in the method of PM_{10} measurement. We calculated the average daily concentration of each pollutant from as many monitors as possible. We included a monitor in the calculation if certain completeness criteria were fulfilled.[6] Despite the completeness criteria, a few missing values remained and were replaced according to the procedure described below.

A missing value on day *i* of year *k* from monitor *j* was replaced by a weighted average of the values of the other monitoring stations as follows:

$$\hat{x}_{ijk} = \bar{x}_{i.k} \times (\bar{x}_{.jk}/\bar{x}_{..k})$$

where $\bar{x}_{i.k}$ is the mean value on day *i* of year *k* among all monitors reporting, $\bar{x}_{.jk}$ is the mean value for monitor *j* in year *k*, and $\bar{x}_{..k}$ is the overall mean level in year *k*.

In ten cities (Athens, Basel, Budapest, Cracow, Erfurt, Geneva, Milano, Rome, Torino, and Zurich), PM_{10} measurements were not available for the whole time period but were estimated using a regression model relating collocated PM_{10} measurements to the BS (for Athens and Cracow) or total suspended particles measurements (for Budapest and Erfurt) or as a percentage of total suspended particles (based on measurements, for the other cities).

Confounders

We used the daily average temperature and relative humidity to control for potential confounding effects of meteorologic variables. All available information on influenza epidemics was recorded, and unusual events (for example, heat waves) were also taken into account. We adjusted for day of the week, national and school holidays, seasonality, and long-term trends. The potential confounding effects of the daily levels of other pollutants (Table 1) were investigated, as described below. The correlation coefficients between PM_{10} and NO_2 ranged from 0.12 to 0.75; between PM_{10} and ozone from -0.38 to $+0.38$; between PM_{10} and SO_2 from 0.14 to 0.78; and between BS and each of these pollutants from 0.11 to 0.65, -0.55 to -0.04, and 0.41 to 0.77, respectively, in the different cities.

Potential Effect Modifiers

Substantial heterogeneity in the estimated effect parameters has been observed previously.[7,16–18] It was therefore important to collect information on several variables hypothesized to be potential effect modifiers. These variables are "city characteristics" (that is, one value per city, which characterizes a particular situation, such as its climate or air pollution sources). The potential effect modifiers for which information was recorded are classified into the following four categories.

(1) Air pollution level and mix. This category includes the average levels of PM (PM_{10} and BS) and that of other pollutants for the whole study period as well as the ratio of PM_{10} and BS to NO_2. The former address the question of whether the effect size depends on the level of exposure *per se* or on the level of exposure to other pollutants. The ratio of PM to NO_2 indicates the extent to which PM comes from traffic, because NO_2 is mainly traffic generated. Therefore, a lower PM/NO_2 ratio reflects a higher proportion of traffic-generated PM.

(2) Climatic variables. It has often been proposed that the air pollution effects estimated are modified by climate,[19] and this theory is supported by seasonal and geographic differences observed previously.[7,16–18] To characterize a city's climate, the mean temperature and relative humidity over the whole study period were recorded. The mean annual daily temperature ranged in our cities from 5.9°C (Helsinki) to 17.8°C (Athens) and the mean relative humidity from 48.9% (Marseilles) to 82.3% (Dublin).

(3) Health status of the population. It is also thought that air pollution affects certain subgroups of the population to a greater extent. Older persons and those suffering from chronic cardiorespiratory disease are obvious candidates. As indicators of their size, the age-adjusted mortality and lung cancer mortality rates for each city's population were used as well as the percentage of persons over 65 years of age and smoking prevalence. The directly standardized annual all-cause mortality rate per 100,000 ranged in our cities from 579 (Lyon) to 1,231 (Lodz). Fifteen cities had values below 800; nine between 800 and 1,000; and five above 1,000. The annual lung cancer mortality rate ranged between 28 and 92 deaths per 100,000 person-years; the proportion of the population >65 years of age between 9% and 21%; and smoking prevalence from 22% to 55%.

(4) Geographic area. It has been observed before[7] (on the basis of fewer cities) that the effect size differed by geographic area. To investigate this further, we classified the cities into three categories: Central-Eastern (which included all cities in the former Communist countries: Budapest, Cracow, Erfurt, Ljubljana, Lodz, Poznan, Prague, Teplice, and Wroclaw), Southern (those with latitude less than 45°: Athens, Barcelona, Bilbao, Madrid, Marseille, Rome, Tel Aviv, and Valencia), and North Western (all other cities). Alternatively we used latitude and longitude.

Analysis

We used a hierarchic modeling approach. First, we fitted regression models in each city separately to allow specific control for seasonal effects, weather, and other potential confounders. We used the results of the individual city analysis in turn in a second-stage analysis to provide overall estimates and to investigate potential effect modifiers.

Individual City Data Analysis

We analyzed the data for each city separately according to a predefined standardized methodology, which resulted in a city-specific model. All data were analyzed in one location (Athens) by three statisticians. We applied generalized additive models (GAM) extending Poisson regression to model the nonlinear effects of the covariates, using a local nonparametric loess smoother to control for seasonal patterns and long-term trends, allowing for overdispersion.[20]

A broad range of smoothing parameters for time removed the basic seasonality from the data. To choose among these, we used diagnostic tools including partial autocorrelation plots and plots of residuals over time to determine the smoothing parameter (that is, the fraction of the data used for smoothing). We had decided in advance that the smoothing window should not be below 2 months to avoid eliminating short-term patterns actually due to the exposure under study. After seasonal and long-term trends were controlled for, we incorporated meteorologic variables into the model. We investigated smoothed functions of the same day and of lags up to 2 days or averaged over 0 to 2 days of daily mean temperature and relative humidity. Same-day values were always included. The inclusion of lagged weather variables and the choice of smoothing parameters for all of the weather variables were done by minimizing Akaike's information criterion.[20] Finally, we added dummy variables to the model to control for day of the week, holidays, or unusual events if necessary.

Information on daily influenza counts was not available for all cities. On the basis of results from a sensitivity analysis (manuscript in preparation), we decided to control for influenza using a dummy variable taking the value 1 for days when the 7-day moving average of the daily respiratory number of deaths was greater than the 90th percentile of its distribution and 0 otherwise. Thus, influenza control was uniform for all cities.

The air pollution variables were put into the model last. Previous results have indicated that in areas with high particle concentrations, log-transformed measurements of particles best represented the mortality-particle relation.[21] To facilitate the second-stage analysis, we decided to use only linear terms, and thus the analysis was restricted to days with BS or PM_{10} concentrations below 150 $\mu g/m^3$. We decided *a priori* to use the average of lags 0 and 1 for BS and PM_{10} measurements. This decision was based on previous studies having shown those lags to be the most relevant.[22] It also avoids potential bias, which could result from selectively reporting those lags associated with the largest effect estimates. We also fitted two-pollutant models to adjust for the confounding effects of SO_2, O_3, and NO_2. Carbon monoxide measurements were not used, because there were many cities with incomplete measurements or without carbon monoxide measurements. If serial correlation remained in the residuals of the final models, autoregressive terms were added. All analyses were done using S-Plus.[23]

Second-Stage Analysis

We applied a second-stage analysis to provide a quantitative summary of all individual city results and to explain heterogeneity, if present. We assumed that the city-specific estimates b_i were normally distributed around an overall estimate, assuming heterogeneity. To examine this heterogeneity we assumed $b_i \sim N(\beta_0 + \gamma z_i, \Omega)$ where β_0 is the mean of the b_i values, z_i is a vector of effect modifiers in city i, γ is the vector of regression coefficients for the effect modifiers, and Ω is the covariance. Such hierarchic models are becoming more common in epidemiology.[24]

To investigate the role of potential effect modifiers, we applied univariate (for one-pollutant models) or multivariate (for two-pollutant models) regression models. We estimated fixed-effects pooled regression coefficients by weighted ecologic regression of city-specific estimates on potential effect modifiers (at city level) with weights inversely proportional to their city-specific variances. If substantial heterogeneity among city results (beyond the variation associated with the effect modifiers) remained, random-effects regression models were applied. In these latter models, it was assumed that the individual coefficients are a sample of independent observations from the normal distribution with mean equal to the random-effects pooled estimate and variance equal to the between-cities variance. We estimated the between-cities variance from the data, using the maximum likelihood method described by Berkey *et al*,[25] and this variance was added to the city-specific variances.

For multivariate second-stage regression models, we applied the method described by Berkey *et al.*[26] In contrast to the usual univariate second-stage regression, in which results from each pollutant are analyzed separately, the multivariate model provides more accurate estimates by incorporating the correlation among pollutants within each city. Specific S-Plus functions (available on request) were written to fit the univariate and multivariate second-stage regression models.

Results

Figure 1 shows the percentage increase in the daily number of deaths associated with 10 $\mu g/m^3$ increase in PM_{10} measurements for each city as well as the pooled estimates. Because there was substantial heterogeneity in the single-city results, pooled estimates using random-effects models are also shown. The estimated increases (per 10 $\mu g/m^3$ increase in PM_{10}) for single cities ranged from -0.6% to 1.5%. The combined increase in the