

Short term effects of air pollution on health: a European approach using epidemiologic time series data: the APHEA protocol

K Katsouyanni, J Schwartz, C Spix, G Touloumi, D Zmirou, A Zanobetti, B Wojtyniak, J M Vonk, A Tobias, A Pönkä, S Medina, L Bachárová, H R Anderson

Department of Hygiene and Epidemiology, University of Athens Medical School, Greece
K Katsouyanni
G Touloumi

Harvard School of Public Health, Boston, USA
J Schwartz

GSF-Forschungszentrum Umwelt und Gesundheit, Germany
C Spix

Faculté de Médecine, Université de Grenoble, France
D Zmirou

Institute of Clinical Physiology, National Research Council, Pisa, Italy
A Zanobetti

National Institute of Hygiene, Warsaw, Poland
B Wojtyniak

Department of Epidemiology and Statistics, University of Groningen, The Netherlands
J M Vonk

Institut Municipal d'Investigació Mèdica, Barcelona, Spain
A Tobias

Helsinki City Center of the Environment, Finland
A Pönkä

Observatoire Régional de la Santé, Paris, France
S Medina

National Center for Health Promotion, Bratislava, Slovakia
L Bachárová

Department of Public Health Sciences, St George's Hospital Medical School, London, UK
H R Anderson

Correspondence to: Dr K Katsouyanni, Department of Hygiene and Epidemiology, University of Athens Medical School, 75, Mikras Asias Street, 115 27 Athens (Goudi), Greece.

Abstract

Background and objectives – Results from several studies over the past five years have shown that the current levels of pollutants in Europe and North America have adverse short term effects on health. The APHEA project aims to quantifying these in Europe, using standardised methodology. The project protocol and analytical methodology are presented here.

Design – Daily time series data were gathered for several air pollutants (sulphur dioxide; particulate matter, measured as total particles or as the particle fraction with an aerodynamic diameter smaller than a certain cut off, or as black smoke; nitrogen dioxide; and ozone) and health outcomes (the total and cause specific number of deaths and emergency hospital admissions). The data included fulfilled the quality criteria set by the APHEA protocol.

Setting – Fifteen European cities from 10 different countries with a total population over 25 million.

Methodology – The APHEA collaborative group decided on a specific methodological procedure to control for confounding effects and evaluate the hypothesis. At the same time there was sufficient flexibility to allow local characteristics to be taken into account. The procedure included modelling of all potential confounding factors (that is, seasonal and long term patterns, meteorological factors, day of the week, holidays, and other unusual events), choosing the “best” air pollution models, and applying diagnostic tools to check the adequacy of the models. The final analysis used autoregressive Poisson models allowing for overdispersion. Effects were reported as relative risks contrasting defined increases in the corresponding pollutant levels. Each participating group applied the analyses to their own data.

Conclusions – This methodology enabled results from many different European settings to be considered collectively. It represented the best available compromise between feasibility, comparability, and local adaptability when using aggregated time series data not

originally collected for the purpose of epidemiological studies.

(*J Epidemiol Comm Health* 1995;50(Suppl 1):S12-S18)

The APHEA project is supported by the European Union Environment 1991-94 Programme. The project must be placed in the context of recent studies investigating the short term adverse health effects of moderate and relatively low air pollution levels which have consistently indicated the existence of effects at levels below the current national and international air quality guidelines.¹⁻⁷ The background and rationale of the study as well as the study areas and air pollution levels have been described in detail elsewhere.⁸

The APHEA project is a multicentre temporal study that uses aggregated data.⁹ Eleven European groups participate, analysing data from 15 European cities, with a total population over 25 000 000. The objectives of the programme are:

- To provide quantitative estimates of the short term health effects (using the total and cause specific daily number of deaths and emergency hospital admissions) of air pollution, taking into consideration interactions between different pollutants and between pollutants and other environmental factors.
- To further develop and standardise the methodology for the detection of short term health effects in the analysis of epidemiological time series data.
- To select and develop a meta-analytic approach for epidemiological time series studies.
- To assess the feasibility of creating a European data base of air pollution measurements and of health indicators recorded on a daily basis. This will allow continuous surveillance of short term effects of air pollution in the future.

The cities involved in the project are (in alphabetical order): Amsterdam, Athens, Barcelona, Bratislava, Cracow, Helsinki, Koeln, Lodz, London, Lyon, Milan, Paris, Poznan, Rotterdam, Wroclaw. There is substantial variability in air pollutants levels, mixtures, and patterns, as well as in the climatic conditions of the 15 APHEA cities. All the cities except one have analysed data over at least five years, typically covering part of the 1980s and beginning of 1990s.⁸

The APHEA collaborators decided on a standardised procedure for testing and quantifying the short term health effects of air pollution and this was applied by each participating group in the analysis of their own data. Thus, exchange of knowledge and expertise and harmonisation of the methodology between the groups were achieved. The data used and their selection criteria are briefly described here, since they have been presented in detail elsewhere,⁸ while emphasis is placed on the statistical evaluation procedures. The analytical methodology is presented from an "applied" point of view; for an in depth statistical review see Schwartz *et al.*¹⁰

The data base

HEALTH OUTCOME DATA

The outcome time series were daily counts of all deaths, or deaths from specific causes (respiratory, cardiovascular, and digestive), or the daily number of emergency hospital admissions from certain causes (total respiratory, asthma, chronic obstructive pulmonary disease, cardiac).

The sources of the mortality data were the national death registration systems for all cities except Athens, where the researchers collected and coded the data. Death registration was complete in all the participating countries. Data on causes of death have problems of incompleteness or erroneous entries and although the extent of these was not studied in each participating city, it was probably relatively small for the diagnostic categories considered in this project.¹¹⁻¹³

Hospital admissions data may have more problems of comparability and uniformity. In most cities¹⁴⁻¹⁶ the data were provided by government operated national registers or municipality operated local registers.¹⁷ In Barcelona, the hospital admissions register was set up and operated by the researchers and has been tested for validity.³ Discharge diagnosis was used in all cases. The coverage was above 95%, and in the cities where this changed over the study period, care was taken to model this trend.¹⁶ Completeness of diagnosis was above 70% for all diagnostic categories and tended to be higher for the categories studied here.^{14,15} Incomplete recording of diagnosis, however, is probably not correlated to air pollution levels on a day to day basis and is therefore unlikely to introduce biases.¹⁰

Errors in the outcome data are probably non-differential with regard to exposure – that is, they do not depend on the air pollution levels. Therefore, if any bias were introduced it would probably tend to decrease the magnitude of the estimated effect parameters (bias toward the null).¹⁸

AIR POLLUTION DATA

The exposure times series data which were analysed concerned sulphur dioxide (SO₂), particulate matter (either total suspended particles (TSP) or particles with aerodynamic diameter smaller than a certain cut off, for example,

PM₁₀, or black smoke (BS)), nitrogen dioxide (NO₂), and ozone (O₃). Daily (24 hour) measurements (from midnight to midnight in all cities except Polish ones) were used for particles, SO₂, and NO₂. For SO₂ and NO₂ the maximum hourly value for each day was also used. For O₃ the maximum one hour and the maximum eight hour daily values were considered.⁸

There was no quality assurance or quality control programme within APHEA to ensure comparability of air pollution measurements. However, all EU countries have quality control programmes in order to conform with EU requirements¹⁹⁻²³ and Finland also has a calibration programme. Furthermore, it has recently been shown that the methods employed in the WHO European Region for measuring SO₂, NO, and NO₂ differ only slightly.²⁴

Admissibility

It was decided to exclude from the analysis measurements done in stations located in limited access highways. Only urban air pollution was studied, so air pollution monitoring sites situated outside urban areas were not used, except for O₃, for which in some instances a "suburban" site was used. All groups used urban background monitoring sites and one site near an urban road, with the exception of the French groups who used only background sites. The arithmetic average of measurements available from all urban stations which fulfilled the completeness criteria (see below) was used for every day. The average from the same stations was used for the whole study period.

All time series studies probably suffer from inaccurate estimation of population exposure, since the same level of exposure is assigned to all members of the population. The number of studies estimating the relationship between individual exposure and the levels measured at fixed monitors are very limited so far.²⁵⁻²⁸ In the APHEA project, however, in order to ensure adequate representation of the population exposure, the number of monitoring sites used for every pollutant was at least equal to three in most cases. The correlation coefficients between the measurements from various monitoring sites in the same city generally ranged from 0.20 to 0.91 for SO₂, 0.16 to 0.93 for black smoke, 0.44 to 0.95 for TSP, 0.20 to 0.84 for NO₂ and 0.40 to 0.97 for O₃.

Completeness

For the calculation of 24 hour NO₂ and SO₂ and maximum one hour NO₂ values, at least 75% of the one hour values on that particular day had to be available. For the maximum one hour O₃ values, 75% of the hourly values from 6 am to 7 pm had to be available, since the maximum O₃ levels always occur during daylight. For the eight hour value of O₃, it was decided to take the 9 am to 5 pm average (since O₃ peaks at or immediately after mid-day and this eight hour average is probably identical or very close to the maximum), and to calculate this, at least six hourly values had to be avail-

able. If a station had more than 25% of the values for the whole period of analysis missing it was excluded. In some centres a station was closed for a long period and some participants used the measurements from another nearby station. In this situation, care was taken not to introduce systematic error, because in some cases a nearby (in geographic terms) station, could give systematically different values (for example if the levels of the substitute station were systematically higher by 25%, they were multiplied by 0.8).

Missing data

In spite of the completeness criteria, there were still a few missing values in the air pollutants time series for some days. It was decided to fill these in since it is desirable to have complete data series for time series analysis. The method was based on estimating the missing value(s) using the available measurements in the other monitoring sites on the same day. If, for example, a value was missing for a specific pollutant from station *i*, the mean of all stations (on days when all measurements were available) was regressed on the measurements of the other sites (used as independent variables) except *i*, adjusting also for season. This was then used to predict the mean daily level when the measurement from *i* was missing, and from this mean to calculate a level for station *i*. This approach is very difficult when there are more than three monitoring sites. In this situation a simpler approach was adopted: the missing value was replaced by the mean level of the remaining stations multiplied by a factor equal to the ratio of the seasonal (three month) mean for the missing station over the corresponding mean from the stations available on that particular day.

DATA ON CONFOUNDING FACTORS

Time series data on daily temperature (°C) and relative humidity (%) were used to control for the potential confounding effects of weather. Information of influenza epidemics and other unusual events (heat wave, strike, etc) was also used when possible, but for the control of other confounding factors (seasonality, calendar effects, long term trends) no collection of data was necessary.

Statistical modelling for control of confounding

STATISTICAL CONSIDERATIONS AND DIAGNOSTIC TOOLS

Time series of the type analysed here are approximately Poisson distributed, overdispersed (that is, the variance is greater than the mean), and usually positively autocorrelated. Overdispersion and serial correlation are usually a result of extraneous factors, such as seasonality and weather, less an intrinsic feature of the daily case counts, and are best taken care of by an appropriate mean model. For details of distributional issues, the relative risk model (Poisson regression), and the principle of in-

cluding potential confounders in the model see Schwartz *et al.*¹⁰ According to the APHEA protocol, each centre controlled individually for each confounder in the a priori list, following the guidelines and procedures set by the protocol. The elements are described in the next paragraph.

For technical reasons, the model building was done using linear regression with log transformed dependent variables. Instead of the Poisson model (where *Y* is the daily number of cases, *X* the independent explanatory variables, *β* the parameter vector, and *E* the statistical expectation)

$$\log(E(Y)) = X\beta$$

we fit

$$E(\log(Y + c)) = X\beta$$

(where *c* is a constant, taking a value >0 when there are 0 counts in the outcome data series, for example, *c*=0.5 or *c*=1) which has a different error structure.¹⁰

Many tools were available for diagnosis and decision making during the process of model building. Some groups used these at each step, and they were used by all groups after finalising the “core” model.

The time series plot of the observed dependent variable

This helped determine whether there was a considerable trend, seasonality, changes in structure (more likely to occur in hospital admissions data than in mortality counts), and shorter unusual periods such as effects of hot or cold spells and epidemics.^{10 29 30}

The predicted time series plot

This depicted the effects of the trend and seasonal model and all other longer term structures. It helped to decide whether a shorter cycle (see below) in the seasonal model represented a genuine change in the seasonal shape. It also helped to determine (together with the observed data plot) whether a certain structure was mostly driven by a short influential phase and, although perhaps statistically significant, was probably not be appropriate for modelling the rest of the series.^{10 29 30}

The residual time series plot

This was most useful in connection with the raw data plot and the predicted series. The goal was to come as close as possible to white noise. Structures such as trend, seasonality, and epidemics visible in the raw series should be adequately described in the predicted series and invisible in the residual series. Structures that showed up in the residual series but were not visible in the original series, pointed to inappropriate or overfitting of one or several confounders. Problems with changes in patterns of longer phases also became visible.^{10 29 30}

Statistical tests; goodness of fit

Tests for improvement of fit (F test, χ^2 test) after the inclusion of one or a group of potential confounder variables were helpful in deciding which terms actually to include in the model. They also helped in deciding between several ways of describing the same phenomenon. The main goal in this type of analysis was not statistical hypothesis testing, it was confounder control. One should not therefore rely exclusively on statistical tests for decision making; the other more descriptive tools presented above are just as important. It should also be taken into account that several correlated variables can jointly lead to good confounder control, while the statistical tests of their parameters are quite misleading. As described in Schwartz *et al.*,¹⁰ R^2 is not an appropriate measure of model fit in the Poisson case. It is, however, possible to check whether the extra-Poisson variation usually observed in the raw data had been completely removed (or removed to a large extent) in the residual series.

Parsimony

Given the usually large number of data points available here, parsimony (that is, the use of the smallest possible number of parameters) was not in the itself important. Unnecessarily large numbers of variables, however, should be avoided. Different descriptions of the same phenomenon sometimes require very different numbers of parameters. For example, a long term trend may be concisely described by a second order polynomial or by a large number of harmonic waves with periods of more than one year (see below). Epidemics that distort the seasonal pattern in some years may require many by-year seasonal terms to describe them, or, alternatively, a small number of (not necessarily rectangular) indicators per epidemic.

Periodogram

The periodogram is part of the spectral decomposition of a time series. Its practical value is that it can be read like an empirical frequency distribution of cyclical patterns of different period lengths in the data. A trend shows up as large values for periods well above one year, seasonality as large values around period 365, and day of week patterns around periods of seven days. After model fitting, the residual periodogram should come as close as possible to white noise, though a formal test of white noise was an insufficient criterion here. Large values in the residual series periodogram after confounder control at periods above 60 days (the limit set for long term pattern removal in this protocol) point to insufficient trend, season, and epidemic control. Large values in the residual periodogram during periods where there were none in the raw data series, point to inappropriate or over correction of at least some potential confounders.²⁹

Partial autocorrelation function (PACF)

The PACF describes the serial correlation of a time series at lags 1, 2, etc, with the value at each lag corrected for the previous lags. For details on the interpretation and treatment of the autocorrelation structure in the regression see Schwartz *et al.*¹⁰ The usually relatively strong positive autocorrelation in the observed health outcome series should be reduced to white noise as far as possible by confounder control. Large positive values at the first lags point to insufficient confounder control. Statistical significance is a helpful but insufficient criterion for what "large" is; this also depends on how large the raw series autocorrelation was. If the first several lags of the residual series are consistently below 0, this points to overcorrection. The PACF of the residuals after completing the confounder model is also helpful in determining the order of the autocorrelation correction in the regression (see below).

Residual-residual plots

When meteorological terms or pollution variables are included in the model, the dose-response curve sometimes requires transformation of the independent variables or a higher order polynomial of them. Smoothed residual-residual plots – that is, using the health outcome variable corrected for all influences except the new one and the new independent variable corrected (pre-filtered) for the same influences – are helpful for determining the shape. Smoothing can be done in a number of ways. Here we chose to plot the means of 20 consecutive days ordered by the pre-filtered new variable. If a pattern does not allow a final decision, but at least restricts the choice to a few alternatives, the decision may then be based on model goodness of fit.

Cross correlations

The delay of the short term effect of meteorology on air pollution can be determined by goodness of fit or with the help of a cross correlation plot. The pre-filtered residuals of the health outcome series and the perhaps transformed influence variable residuals are correlated to a number of delays (for restrictions see below) and the largest correlation is chosen. In series of this type, only one direction of influence is interpretable: if there is any influence, it may only be of the environmental factor on health, not vice versa.

In time series, where the unit of data aggregation is one day, confounding factors are factors likely to vary on a chronological basis over the study period. Such factors are long term trends, cyclical patterns, day of the week and holidays, unusual events (influenza epidemics, strikes), and meteorological variables. In the APHEA data analysis, decisions were taken for the procedures used for the control of all potential confounders.

CONTROL OF CONFOUNDING

Seasonality and other cyclical patterns

Seasonality is an important (probably the most important) confounder when short term effects of air pollution on health are considered. Many health outcomes have pronounced seasonal patterns – for example, mortality counts peak in the winter, a phenomenon which has been observed in places with different latitudes and climates worldwide.³¹

At the same time, the levels of several pollutants exhibit seasonal patterns which, if not controlled for in the analysis, may result in spurious positive (if the pollutants also peak in the winter) or negative (if the pollutant peaks in the summer) associations. The strength and shape of the seasonal pattern depends to a large extent on local conditions.

In the APHEA project it was decided to use a parametric way of controlling for seasonal and other cyclical patterns – that is, through the introduction of sinusoidal terms in the regression model. An annual harmonic (sinusoidal) wave can be described by a function: $\alpha \sin(2\pi t/365) + \beta \cos(2\pi t/365)$, $t=1$ to T , day of study.

While the term $w=2\pi t/365$ determines the frequency or period of the harmonic wave (here: 1 year), α and β determine the amplitude (that is, how strongly the pattern is expressed) and the relationship between α and β determines the phase (that is, where in the course of 1 year the peak and dip of the wave are placed). α and β are determined as regression coefficients. As this pattern is, by definition, symmetric, which may not be a realistic description of the data, further terms with $w=2\pi t/k/365 \cdot 25$, $k=2,3$ etc are included. This can now describe other annual shapes – for example, steep increases and flat declines, or long summers and short winters, or additional peaks and dips beyond the simple summer-winter pattern. For this protocol it was decided to include, at most, terms of the 6th order ($k \leq 6$), as this would pick up two month periods or events of one month's length, which seemed reasonable given that the study aimed at short term effects. If there is reason to believe that the seasonal pattern is different between study years or other sections of the data (with respect to amplitude and/or phase), interaction terms between cyclical terms and indicators for years or sections can be included. Overlaying a biannual cycle ($k=0.5$) can sometimes help in explaining differences in seasonality between years too. Biannual cycles or varying patterns between years can be caused by epidemics, which are better corrected for directly either as (not necessarily rectangular) indicator variables or, where available, as data on case counts. Details of the reasoning behind the inclusion of these terms in the model at all, the advantages and disadvantages of using this method, as well as other methods available to control for seasonal patterns are discussed elsewhere.¹⁰ The maximum order of the seasonal model for an individual city is based on goodness of fit tests as well as on inspection of time series plots and residual periodograms.

The sine and cosine term of a certain period must always be included or excluded simultaneously.

Short term effects of meteorological factors

After the deseasonalisation of the temperature and humidity series, the residuals from the models of temperature and humidity described in the previous section were plotted separately against the deseasonalised mortality residuals. To smooth the curve and reveal any patterns in the association, the days were sorted according to the temperature (or humidity) and the mean of each 20 consecutive days was taken (see above "Residual-residual plots"). The decisions on which temperature and humidity terms to include were based on the shape of this plot, and where it was ambiguous, on model fit differences. Examples of temperature terms which have been used in different locations are as follows:

- Linear term only.¹⁴
- Second order polynomial of temperature or humidity (linear and quadratic term, parabolic dose response curve).^{29 32}
- Double quadratic curve: the inclusion of two variables, one for "cold" taking the values ($A - \text{temperature}$), when temperature was $<A$ (A in °C is the turning point of the double quadratic curve) and 0 otherwise; and one for "hot", taking the value ($\text{temperature} - A$), when temperature $>A$ and 0 otherwise. The level of the turning point in the double quadratic curve (a U shaped curve) depends on the local conditions.³⁰
- Third order polynomial and other variations.¹⁷

In most instances one of the above choices was tested statistically by alternative inclusion in the model. When transformations of temperature and humidity were used, measurements of the same day (lag 0), the previous day (lag 1), and two days before (lag 2) were tested. The same lag was used for temperature and humidity in any single data set. Sometimes inclusion of the temperature-humidity interaction improves the fit.^{14 29} Temperature and humidity variables were kept in the core model irrespective of their statistical significance.

Long term trends

If a systematic increase or decrease in the health outcome was observed over the study period, a variable taking the values 1 to T , where T is the number of days in the study period was included in the model. The square of this variable was also introduced if it improved the model fit. Where there were year to year fluctuations (non-monotonic), dummy variables for the years were included in the model (the number of dummy variables was $N-1$, where N the number of years studied).

Calendar effects

Six dummy variables were included in the model to account for systematic difference in the health outcome in relation to the day of

the week. They were *all* included in the model or *all* excluded, based on model fit. Sometimes the day of week pattern changes in relation to year or season and this was accounted for in the model. Holidays were included as one or several dummy variables, depending on the length of the series and the diversity of their effect.

Influenza epidemics and other events

Appropriate variables were introduced if locally necessary when influenza epidemic data were available or any other unusual event (strike, heat wave, etc)¹⁴ had taken place. For influenza epidemics delayed effects of up to 15 days were tested.

When all core model variables were decided, the time plots of residuals were inspected for any remaining patterns.

Air pollution variables for single pollutant models

Residual-residual plots were examined to decide on the shape of the dose-response curve. For this purpose, the core model was applied with each air pollutant as the dependent variable. Then the residuals of mortality and each pollutant were plotted in order to decide which pollutant transformation to use. Where no decision could be reached on this basis, decisions were made on the basis of model fit. At the next stage, the pollutant variables were included in the model as independent variables. Lags 0 to 3 (0 to 5 for O₃) were checked for one day measurements of each pollutant and cumulative effects by two (lag 0 and 1), three (lag 0, 1 and 2), and four days (lag 0, 1 to 3) (O₃; also five and six days) were also checked alternatively. The observed delay of an effect can be influenced, in addition to biological mechanisms, by a number of local factors such as wind direction, placement of the measuring station, size of the city, "day" definition of pollution measurements, air chemistry, etc. One, one day variable and one indicating cumulative exposure were chosen for each city.

In cleaner settings linear terms for the pollutants generally fitted the data better. When log transformations were the best, these were used to assess the pollutant effects in a city, but in the interest of meta-analysis additional models were fitted with linear pollutant terms, retaining in the analysis only those days when the corresponding pollutant did not exceed 200 µg/m³.³³ Quadratic transformation was also used but specifically for O₃.

Autocorrelation

Remaining autocorrelation was controlled for by using autoregressive error models. Autoregressive terms up to order 4 were tried. Only those that were at least moderately significant (that is, $p < 0.10$) were retained.

Final models

The final models on which the results were based were autoregressive Poisson models that allowed for overdispersion. These models provided relative risk estimates. The SAS software package was used for the above analyses.

Effect modification by season

The possible differential effects of each air pollutant on every health outcome in summer and winter were investigated by introducing an additional dummy variable for season in the model, plus an interaction term of the air pollutant and the season variable. Summer and winter could be differently defined at each centre, depending on the local climatic conditions.

Interaction between pollutants

The same principle used for seasonal interaction was used to investigate possible interactions between pollutants. It was decided to analyse SO₂ effects at different levels of particles (the cut off point for "low" and "high" particle levels was 100 µg/m³ for BS and TSP and 60 µg/m³ for PM₁₀), particle effects at different levels of SO₂ (cut off point for "low" v "high" SO₂: 100 µg/m³), and particle effects at different levels of NO₂ (cut off point for "low" and "high" NO₂ levels were 80 µg/m³ for 24 hour values and at 120 µg/m³ for one hour levels). Thus, in the SO₂ model, for example, a dummy variable was added indicating "low" and "high" particle pollution for each day (usually taking the value 0 for "low" particle days and 1 for "high") plus an interaction term of this dummy variable and the SO₂ variable. In this way, collinearity which presents problems in multi-pollutant models, was avoided.¹⁰

Conclusions

The above procedure was followed by all participants to ensure maximum comparability of the resulting effect estimates. An important advantage of this standardisation is that it enabled us to consider results from many different European settings collectively, without the problem of different methodologies that has beset older studies.^{34,35} Furthermore, the procedure described above allows the flexibility to account for local climate, pollution levels and mixtures, social, and other conditions. It represents the best available compromise between feasibility, comparability, and local adaptability when using aggregated time series data not originally collected for the purpose of epidemiological studies.

The APHEA collaborative group consists of: K Katsouyanni, G Touloumi (Greece; Coordinating centre); D Zmirou, P Ritter, T Barumandzadeh, F Balducci, G Laham (Lyon, France); H E Wichmann, C Spix (Germany); J Sunyer, J Castellsague, M Saez, A Tobias (Spain); J P Schouten, J M Vonk, A C M de Graaf (Netherlands); A Pönkä (Finland); H R Anderson, A Ponce de Leon, J Bower, D Strachan, M Bland (UK); W Dab, P Quenel, S Medina, A Le Tertre, B Thelot, B Festy, Y Le Moulec, C Monteil (Paris, France); B Wojtyniak, T Piekarski (Poland); M A Vigotti, G Rossi, L Bisanti, F Repetto, A Zanobetti (Italy); L Bacharova, K Fandakova (Slovakia).

- 1 Air quality guidelines for Europe World Health Organization, Regional Office for Europe. WHO Regional Publications, European series No 23. Copenhagen: WHO, 1987.
- 2 Kinney PL, Ozkaynak H. Associations of daily mortality and air pollution in Los Angeles county. *Environ Res* 1991; 54:99–120.
- 3 Sunyer J, Anto J, Murillo C, Saez M. Effects of urban air pollution on emergency room admissions for chronic obstructive pulmonary disease. *Am J Epidemiol* 1991;134: 277–86.
- 4 Pope CA, Schwartz J, Ransom MR. Daily mortality and PM₁₀ pollution in Utah Valley. *Arch Env Health* 1992;47: 211–17.
- 5 Schwartz J, Dockery DW. Particulate air pollution and daily mortality in Steubenville, Ohio. *Am J Epidemiol* 1992;135: 12–19.
- 6 Schwartz J. Air pollution and daily mortality in Birmingham, Alabama. *Am J Epidemiol* 1993;137:1136–47.
- 7 Schwartz J. Air pollution and daily mortality: a review and meta-analysis. *Environ Res* 1994;64:36–52.
- 8 Katsouyanni K, Zmirou D, Spix C, et al. Short-term effects of air pollution on health: A European approach using epidemiologic time series data. *Eur Respir J* 1995;8:1030–38.
- 9 Katsouyanni K ed. *Study designs*. Commission of the European Communities, Air Pollution Epidemiology Report Series, no 4. Luxembourg: Office for Official Publications of the European Communities, 1993.
- 10 Schwartz J, Spix C, Touloumi G, et al. Methodological issues in studies of air pollution and daily counts of deaths and hospital admissions. *J Epidemiol Community Health* 1996;50(suppl 1):S3–S11.
- 11 Camilli AE, Robbins DR, Lebowitz MD. Death certificate reporting of confirmed airways obstructive disease. *Am J Epidemiol* 1991;133:795–800.
- 12 Nielsen GP, Bjornson J, Jonasson JG. The accuracy of death certificates. *Virchows Arch* 1991;419:143–6.
- 13 Ashley J, Devis T. Death certification from the point of view of the epidemiologist. *Population Trends* 1992;67:22–28.
- 14 Dab W, Medina S, Quenel P, et al. Short term respiratory health effects of ambient air pollution: results of the APHEA project in Paris. *J Epidemiol Community Health* 1996;50(suppl 1):S42–S46.
- 15 Ponce de Leon A, Anderson HR, Bland JM, Strachan DP. The effects of air pollution on daily hospital admissions for respiratory disease in London between 1987–88 and 1991–92. *J Epidemiol Community Health* 1996;50(suppl 1): S63–S70.
- 16 Schouten JP, Vonk JM, de Graaf A. Short term effects of air pollution on emergency hospital admissions for respiratory disease: results of the APHEA project in two major cities in the Netherlands during 1977–1989. *J Epidemiol Community Health* 1996 50(suppl 1):S22–S29.
- 17 Pönkä A, Virtanen M. Asthma and ambient air pollution in Helsinki. *J Epidemiol Community Health* 1996;50(suppl 1): S59–S62.
- 18 Flegal KM, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. *Am J Epidemiol* 1986;123:736–51.
- 19 Department of the Environment. *Urban air quality in the United Kingdom*. Bradford: Quality of Urban Air Review Group, 1993.
- 20 Conseil des Communautés Européennes Directive du 15/7/1980 concernant des valeurs limitées et des valeurs guides de qualité atmosphérique pour l'anhydride sulfureux et les particules en suspension (80/779/CEE). *Journal Officiel des Communautés Européennes* 1980;L229: 30–48.
- 21 Conseil des Communautés Européennes Directive du 7/3/1985 concernant les normes de qualité de l'air pour le dioxyde d'azote (85/203/CEE). *Journal Officiel des Communautés Européennes* 1985;L87:1–7.
- 22 Conseil des Communautés Européennes Directive 21/6/1986 modifiant la directive 80/779/CEE concernant des valeurs limitées et des valeurs guides de qualité atmosphérique pour l'anhydride sulfureux et des particules en suspension (89/427/CEE). *Journal Officiel des Communautés Européennes* 1989;L201:53–55.
- 23 Conseil des Communautés Européennes Directive 92/72/CEE du conseil du 21/9/92 concernant la pollution de l'air par ozone. *Journal Officiel des Communautés Européennes* 1993;L297:1–7.
- 24 Mucke HG, Manns H, Turowski E, Nitz G. *European intercomparison workshop on air quality monitoring – measuring of SO₂, NO and NO₂*. Vol 1 Berlin: WHO Collaborating Centre for Air Quality Management and Air Pollution Control at the Institute for Water, Soil and Air Hygiene, Federal Environmental Agency. Report 7,1995.
- 25 Environmental Protection Agency. *The total exposure assessment methodology (TEAM) study*. Summary and analysis Vol 1. Washington DC: Office of Research and Development, EPA, 1987.
- 26 Spengler JD, Treitman RD, Tosteson TD, Mage DT, Soczek ML. Personal exposures to respirable particulates and implications for air pollution epidemiology. *Environmental Science and Technology* 1995;19(8):700–7.
- 27 Wallace LA, Pellizzari ED, Hartwell TD, Sparacino CM, Sheldon LS, Zelon H. Personal exposures, indoor-outdoor relationships and breath levels of toxic air pollutants measured for 355 persons in New Jersey. *Atmospheric Environment* 1985;19(10):1651–61.
- 28 Moschandreas DJ. Exposure to pollutants and daily time budgets of people. *Bull NY Acad Med* 1981;57(10):845–59.
- 29 Spix C, Wichmann HE. Daily mortality and air pollutants: findings from Köln, Germany. *J Epidemiol Community Health* 1986;50(suppl 1):S52–S58.
- 30 Touloumi G, Samoli E, Katsouyanni K. Daily mortality and air pollution from particulate matter, sulphur dioxide, and carbon monoxide in Athens, Greece, 1987–91. A time-series analysis within the APHEA project. *J Epidemiol Community Health* 1996;50(suppl 1):S47–S51.
- 31 Lipfert FW. *Air pollution and community health. A critical review and data source book*. New York: Van Nostrand Reinhold, 1994.
- 32 Sunyer J, Castellsague J, Saez M, Tobias A, Anto JM. Air pollution and mortality in Barcelona. *J Epidemiol Community Health* 1986;50(suppl 1):00–00.
- 33 Wojtyniak B, Piekarski T. Short term effects of air pollution on mortality in Polish urban populations – what is the different? *J Epidemiol Community Health* 1996;50(suppl 1):S36–S41.
- 34 Hatzakis A, Katsouyanni K, Kalandidi A, Day N, Tri-chopoulos D. Short-term effects of air pollution on mortality in Athens. *Int J Epidemiol* 1986;15:73–81.
- 35 Derrienic F, Richardson S, Mollie A, Lellouch J. Short-term effects of sulphur dioxide pollution on mortality in two French cities. *Int J Epidemiol* 1989;18:186–97.